2016

Molecular Simulation Studies of Hydrophobins near Gas, Oil and Water Interfaces

Yuwu Chen
Louisiana State University and Agricultural and Mechanical College, chenyuetian1984@gmail.com

MOLECULAR SIMULATION STUDIES OF HYDROPHOBINS NEAR GAS, OIL AND WATER INTERFACES

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Gordon A. and Mary Cain Department of Chemical Engineering

by
Yuwu Chen
B.S., Nanjing University of Science and Technology, 2006
M.S., Nanjing University of Science and Technology, 2008
M.S., The University of Toledo, 2011
M.S., Louisiana State University, 2016
May 2017

Acknowledgements

This project was partially supported by a grant from the Gulf of Mexico Research Initiative (GoMRI), as part of the Consortium for the Molecular Engineering of Dispersant Systems (C-MEDS, http://dispersant.tulane.edu/). High performance computational resources for this research were provided by High Performance Computing at Louisiana State University (http://www.hpc.lsu.edu), and the Louisiana Optical Network Initiative (http://www.loni.org). I am grateful to all colleagues in Hung group and Russo group for their suggestions. Positive discussions and encouragement by the following scientists/scholars are deeply appreciated: Erik G. Marklund, Uppsala University, Sweden; Djurre H. de Jong, University of Groningen, The Netherlands; Ann H Kwan, University of Sydney, Australia; Yuan Fang and Kyle Huston, University of Michigan; and Zenghui Zhang, Michal Brylinski, Xiaoxia He and I-Ting Liu, Louisiana State University.

Table of Contents

Abstract

This dissertation contains fundamental, classical molecular simulation studies of the properties of hydrophobins (a unique family of surface-active proteins produced by filamentous fungi in soil) near interfaces involving gas, oil, water, organic solvents and polymers. These studies are relevant to possibly using hydrophobins as natural oil dispersants and in processing of polymers. Preliminary studies by the Russo group suggest that these small surface-active proteins can encapsulate oil, gases and polymers in *cylindrical* structures. We have performed classic molecular dynamics (MD) simulation and potential of mean force (PMF) calculations of a class I hydrophobin EAS using both all-atom and coarse-grained representations. The interfacial properties of these hydrophobins at gas/water and oil/water interfaces were probed, and our simulation results qualitatively agree with experimental observations. According to the PMF calculation results, EAS molecule was likely to stay at the hydrophobic/hydrophilic interface, and the adsorption behavior of EAS at the interface was strong and irreversible. We performed MD simulations using Martini coarse-grained (CG) models to gain insight into the stability of nm-sized 'blobs' formed by the assembly of hydrophobin around oil. Finally, we developed a Martini model for Poly(γ-stearyl α,L-glutamate), PSLG, and used it to perform MD simulations of PSLG molecules near a hydrophobin-coated hydrophilic-hydrophobic interface. These simulations suggest that initial alignment of the PSLG chains, PSLG concentration and solvent type may affect the final alignment of the PSLG chains.

Chapter 1. Introduction and General Objectives

Hydrophobins are a unique protein family with distinctive surface properties from fungi. In nature, a classical example of fungus is the common mushroom (Figure 1), *Agaricus bisporus*, which is one of the most commonly and widely consumed mushrooms of normal diet in the world. The growing substrate hyphae of common mushroom specifically secrets a hydrophobin called ABH3 [1]. In fact, all hydrophobins exhibit high surface activity, serving as coating and protective agent, in adhesion, surface modification, or other types of function that require surfactant-like properties. The surface-active properties of hydrophobins facilitate the formation of essential aerial structures such as hyphae or spores [2, 3].

Figure 1. Common mushroom, *Agaricus bisporus,* is a fungal secreting hydrophobin ABH3.

1.1 Preliminaries

Before these surface-active proteins were formally named “hydrophobins”, some of these proteins were isolated from various sources of fungi, and its physical properties such as surface activity were systematically studied. For example, according to the works of Russo *et al*. in the 80s [4-6], *Cerato ulmin* (CU), which is produced by the filamentous fungus *Ophiostoma ulmi* and has been implicated in Dutch elm disease [7-9], showed surface activity at concentrations as low as 30 $ng \cdot mL^{-1}$. A number of physico-chemical measurements including light scattering,

nuclear magnetic resonance (NMR) and light microscope on CU in pure water, ethanol/water and gas/water had been carried out respectively by Russo *et al*. By gently rocking the CU solution sample by hand, “rods” and “fibrils” can be easily produced, as displayed in Figure 2 [6]. Their light scattering data clearly demonstrated that hydrophobin CU has extremely low solubility in 70% ethanol/water solution. However, even at such low concentration, hydrophobin can stabilize air bubbles in both spheroidal and *cylindrical* morphology. The relationship between these observations and elm trees infected with the CU producing fungus *Ceratocystis ulmi* suggests a novel delivery mechanism whereby CU at exceedingly low concentrations may be concentrated and transported to intercellular openings, thereby blocking the vascular system of elm trees [5].

Figure 2. Optical micrograph of a 2×10^{-4} g·cm^{-3} CU solution after gentle rocking, showing the presence of "rods" and "fibrils." Distance marker is 50 µm. The figure is a reprint of Ref [6] with permission from NRC Research Press, the publisher of Canadian Journal of Botany.

According to several reviews [10-14], hydrophobins were discovered by Schuren and Wessels around 1990 as the putative products of members of a gene family highly expressed during formation of aerial hyphae and fruit bodies in the homobasidiomycete *Schizophyllum commune* (Sc). This is because they identified the nucleotide sequences of the Sc3 and Sc4 genes of the Sc, and the deduced amino acid sequences. The gene sequence for Sc1 discovered by them earlier was corrected at the same time. Schuren and Wessels also verified that all of the three independently isolated genes (Sc1, Sc2 and Sc4) had similar structures and nucleotide sequences of their coding regions, and the proteins were associated with the genes. Based on the classification methods of amino acid sequences and unique structure of hydrophobin, other scholars were then able to identify their surface active proteins found in the early years as hydrophobins. Take the previous introduced CU involved in Dutch elm disease as an example, it is identified as a hydrophobin by Stringer and Timberlake after making necessary comparison of amino acid alignment with Sc1, Sc2, Sc4 and other hydrophobins [15-17].

1.2 Structure of hydrophobins

Therefore, how hydrophobins can be distinguished from other proteins? Compared to most of proteins, hydrophobins have much lower molecular weights and shorter amino acid chains. Among 50 members of the hydrophobin family that are known to us, the typical molar masses are 10 kDa. A hydrophobin molecule usually has 80-120 amino acid resides. In contrast, yeast protein molecules have an average of 466 amino acid resides and a molar masses of 53 kDa [18]. Nevertheless, the most appropriate method to distinguish the hydrophobin is through the examination of their unique structure.

The initial comparison of sequences among various hydrophobins showed that the sequences of each kind of hydrophobin were diverse. Although hydrophobins have low amino

acid sequence homology, all of them feature a particular conserved repeat of 8 cysteine (Cys) groups in a special pattern. Figure 3 displays the alignment of the hydrophobins. For each

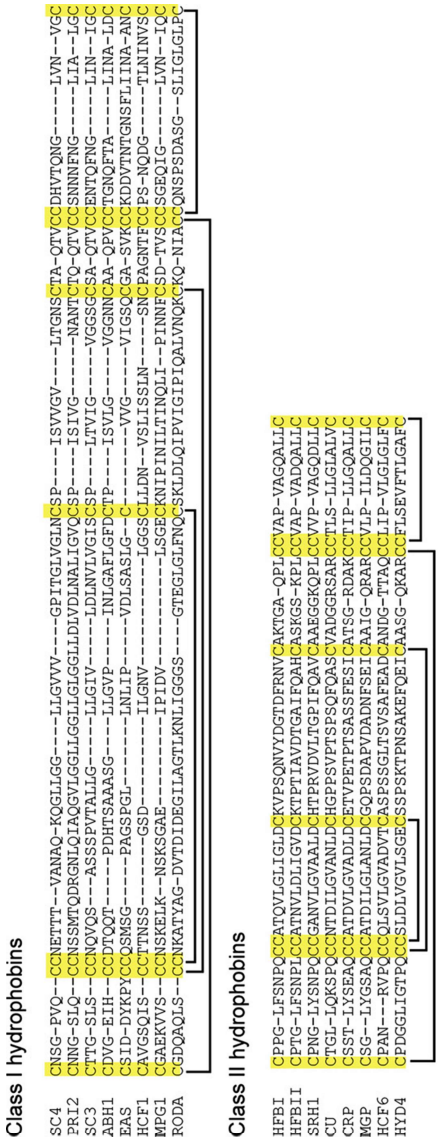


Figure 3. Amino acid sequence (deduced) comparison of class I and II hydrophobins. Reprint of the work of Sunde *et al*. in Ref [14] with permission from Elsevier, the publisher of Micron. Only amino acid residues between the first and last Cys residues are shown as the sequence variance at the termini interferes the structure illustration. The Cys residues are highlighted in yellow. The listed hydrophobins and the works determining their structural gene and amino acid sequence are: SC4 of *S. commune* [10], PRI2 of *A. aegerite* [19], SC3 of *S. commune* [20], ABH1 of *A. bisporus* [1], EAS of *N. crassa* [21], HCF1 of *C. fulvum* [22], MPG1 of *M. grisea* [23], RODA of *A. nidulans* [16], HFB1 of *T. reesei* [24], HFBII of *T. reesei* [25], SRH1 of *T. harzianum* [25], CU of *O. ulmi* [17], CRP of *C. parasitica* [26], MGP of *M. grisea* (accession O94196, unpublished data), HCF6 of *C. fulvum* [27] and HYD4 of *G. moniliformis* (accession Q6YF29, unpublished data).

hydrophobin, four pairs of disulfide bridges are formed by 8 Cys, as shown in black elbow connectors, and the connection sequence is the same: the first Cys residue is connected to the sixth Cys residue; the second is connected to the fifth; the third is connected to the fourth; and the seventh is connected to the eighth. All of the hydrophobins are found in this special pattern, and then they can be defined as "hydrophobin" accordingly, even their amino acid residue sequence are extremely diverse.

Meanwhile, the second and third Cys residues follow each other in the sequence of the main chain, binding by the peptide bond. A pair is therefore formed, and the pair formation is the same as the sixth and seventh Cys residues. Linder in a review [13] stated this pattern of hydrophobin (separated, pair, separated, separated, separated, pair, separated) had a remarkable symmetry and could be easily recognized in a primary sequence.

In addition to the two characteristics of Cys connection mentioned above, there is another characteristic of hydrophobins that is related to their amino acid sequence: all hydrophobins have similar arrangement of hydrophobic domains. For instance, Stringer and Timberlake [15] compared the hydrophobicity plot of CU [28] to those of RODA [16] and SSGA [29], and demonstrated the array of hydrophobic domains in CU is similar to arrays in other fungal hydrophobins. Wessels [12] made a comparison of hydropathy patterns of Sc3, Sc4 and Sc1 to other other six kinds of hydrophobin in 1994, stating if aligned by the eight Cys residues, the compared hydrophobins have striking similarity of the hydropathy patterns.

Based on the secondary structure and aqueous solubility difference, hydrophobins can be divided into two classes: class I and class II [12]. The former, most of which do not have α-structure, are more hydrophobic than class II. Figure 4 reveals the secondary structures of class I hydrophobin EAS in solution and class II hydrophobin HFBII in crystal status. Their secondary

structures were respectively determined by Kwan *et al*. [30] and Hakanpaa *et al*. [31]. An alignment is performed by comparing structures with the program called Structural Alignment of Multiple Proteins (STAMP), algorithm of which minimizes the alpha carbon distance between aligned residues of each molecule by applying globally optimal rigid-body rotations and translations [32]. Once STAMP finished the alignment of two hydrophobin molecules (EAS and HFBII), it overlaid the two structures. As seen in the Figure 4, there are similarities of β barrel structures between EAS and HFBII, but the later has a small segment of α helix.

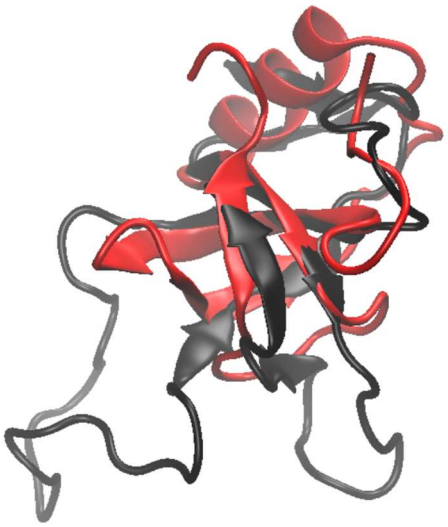


Figure 4. Secondary structures of class I hydrophobin EAS and class II hydrophobin HFBII (overlapped) Red: HFBII; black: EAS. The second structures are represented by the NewCartoon method in VMD [33]. Arrows are assigned to represent β structures.

On the other hand, the secondary structure of EAS in solution, which was determined by NMR measurement, not only has four-stranded β-barrel (K15–C19, V46–V47, S57–K62, and N79–V81), but also includes two-stranded antiparallel β-sheet (T2–T3 and Q53–C54). A simple molecular dynamics (MD) simulation was carried out to test the stability of the β structure of EAS in vacuum or pure water. Then secondary structure evolution was determined by Define Secondary Structure of Proteins (DSSP) program [34, 35]. The results, as shown in Figure 5,

indicate that at 1 bar and 298K, most of the β structures are stable over the 10ns simulation, but one of the β-barrel (V46–V47) and both of the β-sheet (T2–T3 and Q53–C54) had morphology change. Those "disappeared" β structures are notably short. However, the β structure investigation carried out by MD simulation was lack of experimental approval, and may be as a result of immature simulation design or inappropriate use of DSSP program. Stated in another way, though the simulation parameters or the DSSP algorithm might not be well orientated for the EAS in water or vacuum, the results of simulation suggested that the β structures with longer amino acid residues are more likely to be stable in pure water or vacuum.

Figure 5. MD simulation of EAS in water (middle strip) and vacuum (bottom strip). The β structures are shown in red; the disordered parts are shown in other colors. The final secondary structure is determined by Define Secondary Structure of Proteins (DSSP) program [34, 35]. The simulation ran for 10 ns. The top strip is the initial structure before the MD simulation, provided by Kwan *et al.* [30]. NPT simulation (1 bar and 298K) is for the EAS in water while NVT simulation (298K) is for the EAS in vacuum. More details about the force field and models for MD simulation will be introduced in the next chapter.

Besides ordered α helix and/or β structures, Figure 4 also displays some longer disordered parts in EAS rather than in HFBII. There are two determined disordered loops M22–S42 and V65–F72. According to a test of root mean square fluctuation (RMSF) of EAS residue with respect to the beginning structure, as compared to other parts, these two disordered loops fluctuate significantly more intense (Figure 6). Two pairs of the four disulfide bridges (19–45 and 61–80) are inside the β barrel center of molecule, whereas the other two (9–60 and 18–54)

connect the outside surface of the β barrel to a β sheet and a nearest loop. More detailed structure of EAS has been well introduced in the works of Kwan *et al.* [14, 30, 36].

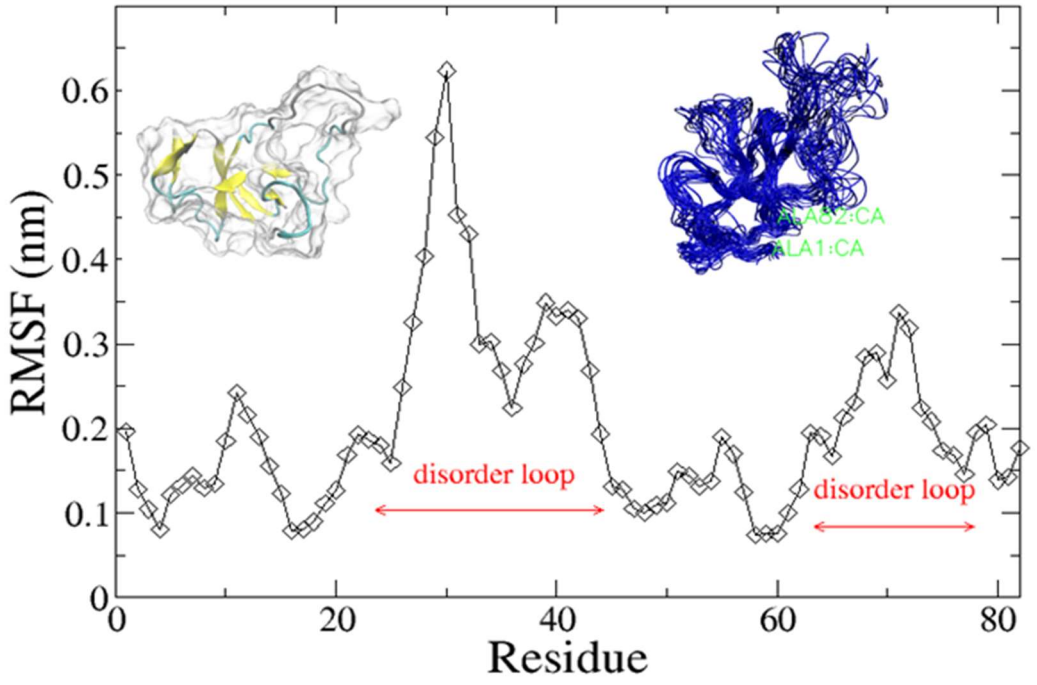


Figure 6. Root mean square fluctuation (RMSF) of EAS residues with respect to the beginning structure. The test was obtained from a 10 ns MD simulation of EAS in water. Left subfigure is the NMR structure of EAS. Blue spaghettis in right subfigure are the superposition of 20 structures sampled every 0.5 ns from a 10 ns simulation of EAS in water. Right subfigure is a reproduction of Ref [36], with permission from Elsevier, the publisher of Journal of Molecular Biology.

The comparison of the secondary structure between EAS and HFBII introduced above also generally applies for other class I and class II hydrophobins. The structure of hydrophobins can explain their solution properties.

1.3 Surface-active and amphipathic properties of hydrophobins

Hydrophobins are products of fungi, but so far the class I hydrophobins can be found in both basidomycetes and ascomycetes of fungi, whereas class II hydrophobins can be found only in ascomycetes of fungi. Yet, fungi produce hydrophobins, expecting them to change the environment surrounding the fungi, because fungi acquire nutrients in a different way than

members of the plant and animal kingdoms. As we know, the plant can directly make use of sun light, turning radiant energy to chemical energy; the animal can eat and break down food in the alimentary canal into substances that can be absorbed and used by the body. On the contrary, fungi are essentially embedded in their host for food with aqueous environment [37]. Lacking mobility, if fungi need to escape from one environment to the other for some purposes such as spreading spores, hydrophobin can help.

So do hydrophobins have some unique properties that can assist fungi? The answer is yes. All of the hydrophobins discovered so far, according to many published works since 1980s [3, 6, 11, 20, 30, 38-47], are surface active and can self-assembly at hydrophilic-hydrophobic interfaces, forming amphipathic surface membranes. Acting as bio-surfactant, hydrophobins can greatly reduce the interfacial tension when self-assembling at those interfaces [31, 48-52]. The surface activity required for self-assembly can begins at concentrations of hydrophobin as low as 50 $\mu g \cdot mL^{-1}$, and the surface tension of water can be reduced from 72 mN/m to 24 mN/m [3, 53]. With the help of hydrophobins, the formation of essential aerial structures of fungi such as hyphae, spores, and fruiting bodies can be facilitated [2].

Wösten in a classic review of hydrophobin [3] offered an interesting and vivid example to explain “how a fungus escapes the water to grow into the air”. The quoted part is the title of a journal article [53], in which Wösten *et al.* originally showed the model illustrating the hydrophobin SC3’s function in the escape of fungi *S. commune* hypha from aqueous environment to aerial environment. Wösten then, in the later review article, made the model more detailed.

As seen in Figure 7, when an aerial hypha of *S. commune* approaches near the water side of the gas/water interface, it begins to secret hydrophobin monomers that can self-assembly at

the interface even at an extremely low concentration. Once an amphiphilic membrane is formed, covering the nearby water surface, the surface tension of water is reduced so the hyphae's

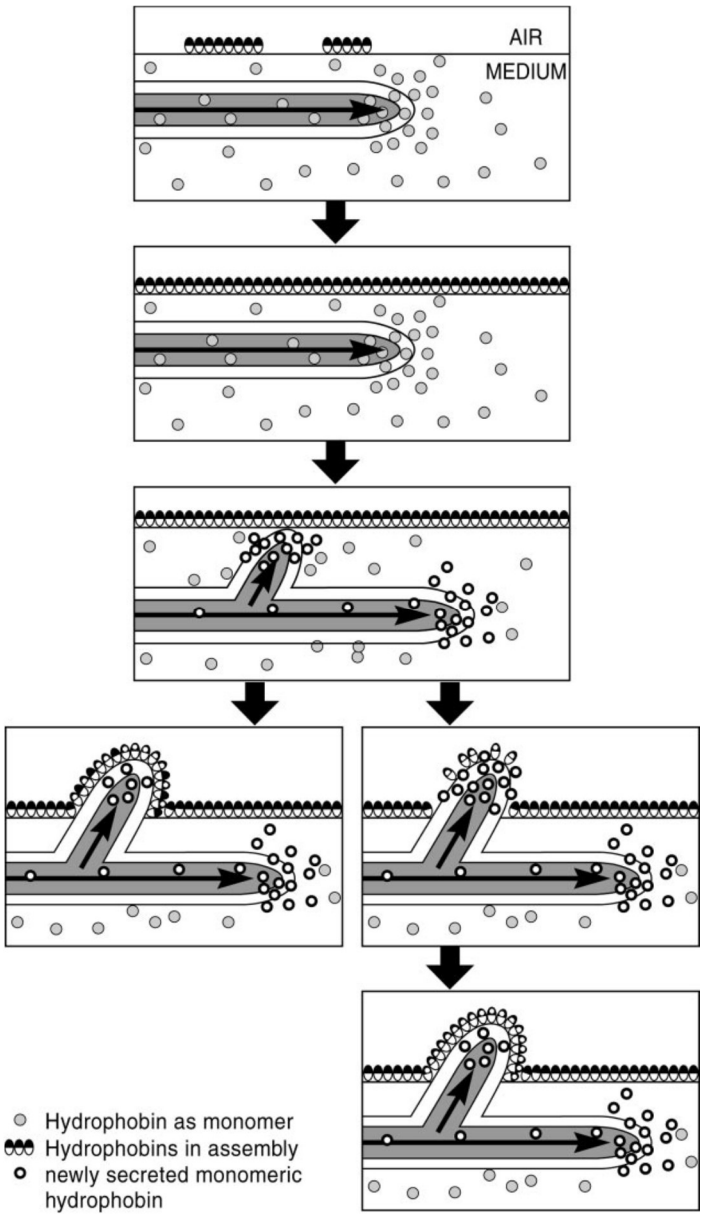


Figure 7. Model for the formation of aerial hyphae in *S. commune*. The figure is a reprint of Ref [3]. According to the Annual Review's publisher, the permission of reprinting this material in a thesis / dissertation is not required. The formation process is in the order of top-bottom. The model contains two hypotheses about how a hypha approaches the gas/water interface and is confronted with the amphiphilic hydrophobin membrane: (1) bottom left, the hypha reaches and extends the hydrophobin membrane with newly secreted hydrophobin monomers, but does not ruptures the membrane and remains in the aqueous environment; (2) bottom right, the hypha physically punctured the hydrophobin membrane and reaches to the air. The newly secreted hydrophobin monomers then cover the hyphae.

expanding to the gas/water interface becomes more easily. The model contains two hypotheses about how a hypha is confronted with the amphiphilic hydrophobin membrane. The first hypothesis states that the hydrophobin membrane is not physically punctured but extended by the force exerted by the turgor pressure of the hypha contacting with the membrane, while more newly secreted SC3 monomers are served as the extended area of the membrane (intercalation). As a result, the hypha would never leave the water phase but is covered by a layer of hydrophobin. As an alternative, the hypha physically punctured the hydrophobin membrane and reaches to the air. The newly secreted hydrophobin monomers can then cover the entire hyphae at the gas/water interface. The biological role of hydrophobin at the gas/water interface is not fully understood, but no matter which model dominates during the formation of aerial hyphae, hypha do have the capacity to break the amphiphilic hydrophobin membrane [53].

Besides discussing the aerial transport mechanism when hydrophobin is assembled at the gas/water interface, scientists are also interested in the properties of surface membranes during hydrophobin's self-assembling. Hydrophobin has been considered as nature's Janus particles [13] because one nominally flat face of the globular protein is coated with hydrophobic groups, while the other face presents hydrophilic moieties.

Hydrophobins are divided into two classes not only based on their secondary structure difference, but also according to the solubility in alcohol and surfactant solutions. When self-assembling, both classes tend to form rodlike structures (Figure 8). Class I hydrophobins are highly insoluble proteins, their self-assemble membrane has rodlet appearance, and cannot be dissolved by organic solvents, hot detergent or treatment with alkali. The rodlets formed by EAS, as shown in Figure 8, was scanned by atomic force microscopic device from Sunde *et al.*[14]. Vocht et al [20] used scanning force microscopy and estimated the SC3 rodlets with an average

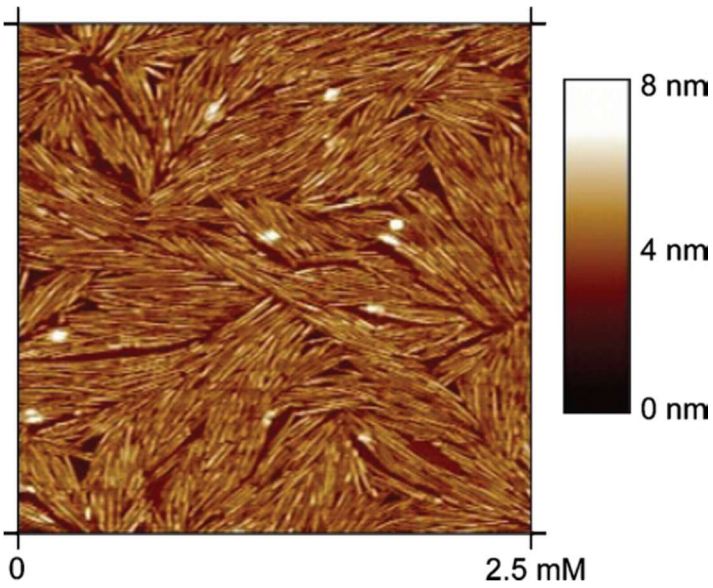


Figure 8. Atomic force micrograph of rodlets formed by purified EAS. Reprint of the work of Sunde *et al.* in Ref [14]with permission from Elsevier, the publisher of Micron.

diameter of 9–15 nm and thick of 7–8 nm. However, the rest composition of the rodlet other than hydrophobin was still unknown. The rodlet membrane can only be dissolved in certain strong acids such as 100% trifluoroacetic acid [43]. On the other hand, the self-assemble membrane of class II hydrophobins does not have rodlet appearance and can dissociate again reversibly by adding ethanol or sodium dodecyl sulfate [46, 54].

Hydrophobins can adhere various surfaces and affect their physico-chemical properties even at low surface coverage concentration. For instance, binding at Teflon particles, the originally hydrophobic surface of Teflon® was wetted-in by hydrophobins [45, 49, 55, 56]. The Teflon® particles then could be dispersed more conveniently in the aqueous solution. The hydrophobins in two classes do not necessarily display the same characteristics on all surfaces. Lumsdon *et al.* [49]shows the class I hydrophobin SC3 did not stabilize Kevlar® nonopulep dispersion, whereas the class II hydrophobin HFBII delayed the sedimentation process of Kevlar®. The results indicated that the adhesion made by class II hydrophobin were stronger. However, on the contrary, Askolin *et al.* [56] tested the binding strength of SC3, HFBI and

HFBII to the hydrophobic Teflon® surface, and found SC3 and HFBI are binding stronger than HFBII. Therefore, different classes of hydrophobin differ in surface adhesion.

1.4 Potential uses

For practical applications, class I hydrophobins are considered more suitable by Subkowski *et al.* [57] because they are generally more stable to high temperature, detergent and solvents than class II. Nevertheless, class II hydrophobins are easier to be isolated than the class I [13]. Despite the subtle difference on structure and hydrophobicity, researchers have proved both classes of hydrophobin can self-assemble at a variety of hydrophobic/hydrophilic interfaces [3, 36, 39, 40, 51, 55, 58-68]. Ignited by hydrophobins' excellent surface properties, researchers have realized various practical applications of these proteins [13, 69]. In this section, several of application potentials of hydrophobins that may have further developments have been reviewed.

Hydrophobin is used as beverage gushing indicator as it can self-assemble at the gas/water interface and forms bubbles (foam) in beer. Gushing is the occurrence of large amount of foam when opening the beer container without agitation. Sarlin *et al.* isolated hydrophobin ELISA from the strains of *Fusarium* and other two fungi, using them to predict the beer gushing activity of malt [70], and found the gushing could happen at a very low hydrophobin concentration of 0.003 ppm. From the work of Sarlin *et al.*, there was a correlation between the hydrophobin concentration and the beer gushing probability of malt. In contrast, there was no correlation observed between the hydrophobin concentration and the *Fusarium* mycotoxin deoxynivalenol level, nor between the later and the gushing potential of the malt studied.

The gushing detection with hydrophobins has been reported by Stübner *et al*. as well [71]. The surface activity of a class II hydrophobin FcHyd5p and its contribution in the gushing was studied, suggesting just a tiny amount (215 μg) of FcHyd5p can induce spontaneous overfoaming

of carbonated liquids. However, biotechnologists have actively applied hydrophobins' self-assembly capability, treating hydrophobins more than just a bio-chemical indicator.

Hydrophobin can be used as stabilizers of emulsions in creams and ointments of personal care [11]. This is because its excellent capabilities of self-assembly at the hydrophobic/hydrophilic interfaces can stabilize emulsions. Thus, there is a great potential for hydrophobins enter the personal care industry, especially in the hair care, which is a highly profitable area. The prospective of large profit may, on the other hand, stimulate the research and development (R&D) intensity and production level of hydrophobins. Hektor in the review of hydrophobin [69] states the self-assembly property of hydrophobin might also be beneficial for applications in the pharmaceutical and food industries, which both require stable emulsions for certain formulations and ingredients.

Hydrophobin has been applied as immobilizer of enzymes in biosensors provided by Palomo *et al.* [72]. In their experiment, Hydrophobin of *Pleurotus ostreatus* was attached on a hydrophilic matrix (agarose) with its hydrophilic part, forming a support for the lipases from *Candida antarct*ica, *Humicola lanuginosa*, and *Pseudomonas flourescens*. The hydrophobic side of the hydrophobin layer successfully immobilized the lipases. After the noncovalent immobilization of lipases to the hydrophobin, it is not necessary to apply detergents or cosolvents to disperse the lipases. The studies of lipase activity was facilitated in its dispersed status.

Hydrophobin can be used as air dispersant in the ice cream, and its relative application prospective has been discussed in a report designing multiscale structures for desired properties of ice cream provided by Crilly *et al*. [73]. "Ice cream is a complex multiphase structure consisting of ice, air, and fat as dispersed phases at a range of different length-scales". However,

the air and ice phases are intrinsically unstable stable, and will coarsen overtime, reducing "creaminess/smoothness, colder eating, poorer visual appeal, and rapid melting". Hydrophobin, which is comparable to the ice structuring proteins, can help ice cream achieve a finer air phase microstructure, preventing large air cells that can cause the freezer of the ice cream. Therefore, Crilly *et al*. realized that in fact the nature has already provided hydrophobins as a direct and effective solution for the microstructure control of ice cream.

Using hydrophobins as dispersant is not confined in the small areas involving organics, such as beer, enzyme, and ice cream introduced above. The water body is a much larger stage for hydrophobins, where biotechnologists have found their potential for numerous applications. Wang et al applied 0.1 $mg \cdot ml^{-1}$ class I hydrophobin from *Grifola frondosa* to disperse multi-walled carbon nanotubes in water [74]. With 30-min sonication, the carbon nanotubes can be effectively on-covalently bonded by hydrophobin through hydrophobic interaction, being rendered hydrophilic.

In brief, the hydrophobins' potentiality for use is promising. The nature is generous for providing hydrophobins, which are bacterially expressed in such large quantities (nearly 1 gram per liter) that commercial applications can be considered [3, 13, 57, 69, 74, 75]. BASF has also succeeded in synthesizing the hydrophobin on an industrial scale [76]. The commercial production level of hydrophobins could be higher if more successful application can be successfully carried out.

1.5 General objectives

There are some classical reviews available to further discuss the basic properties of hydrophobins [2, 3, 13, 14, 69]. We are particularly interested in exploring the feasibility of using hydrophobins as 'natural' oil dispersants under seawater conditions. During the Gulf of

Mexico 2010 oil spill, over one million gallons of commercial synthetic surfactants were released into the Gulf of Mexico to disperse the oil spilled. On the other hand, hydrophobins are abundant in nature and of interest for medical and technical applications [77]. For example, hydrophobin ABH3 from common mushroom have been consumed in large amount by people. Back to the ocean, over years and years every sea creature has been exposed to hydrophobins, so toxicity of hydrophobin should be much reduced compared to synthetic surfactants. The goal of our project is to see whether hydrophobin can be used as a substitution to synthetic surfactants.

A previous study from the Russo group with the class II hydrophobin Cerato-ulmin (CU) suggests that these proteins can efficiently encapsulate oil (to form "blobs") and air (to form "bubbles") in *cylindrical* structures [4-6]. Besides their abundance in nature, hydrophobins have been successfully biosynthesized on an industrial scale [76]. Because of their ease of (biosynthetic) manufacture, it is possible to imagine stockpiling hydrophobins for potential use as 'natural' oi spill dispersants. However, a fundamental understanding of the interfacial properties of hydrophobins near oil, gases and water will be crucial to determine whether these proteins can be used as efficient oil/gas dispersants.

In particular, at first the properties of hydrophobins near relevant interfaces (e.g., gas/water and oil/water) need to be further investigated. For example, the class II hydrophobin CU is known to produce an especially wide variety of very unusual structures in water. With slightly rocking or agitation, *cylindrical* air bubbles can be generated in CU solution, which is contrasted to the energy minimization principle. As we know, most common bubbles and blobs are spherical to energy minimize their status. The unusual shape is stabilized by a thin, solid amphiphilic membrane of the hydrophobin. Once a *cylindrical* bubble (blob) is formed, the hydrophobic/hydrophilic interface is formed, understanding what is happening at the interface is

crucial. Progress is expected on understanding the intermediate structures that connect a small protein and bubbles (blobs) of unusual shape. The simulations is to provide molecular-level feedback. We first perform molecular dynamics (MD) simulations for a class I hydrophobin, EAS, near gas/water and oil/water interfaces. Interfacial properties (free energies, density profiles, radius of gyration of the hydrophobins) that are relevant for the possible use of hydrophobins as 'natural' oil dispersants were probed in our simulations using both all-atom (AA) and coarse-grained (CG) models.

Next, bubbles and blobs exist over an enormously wide range of sizes, requiring a comparably broad set of physical polymer/colloid/particle characterization with MD simulations. With conventional experimental tools, the minimal size of hydrophobin bubbles and blobs in terms of radius of gyration (Rg) have been determined (~10nm), but the shape of the bubbles and blobs at such size scale is undeterminable by experiment. Therefore, we present CG MD simulation details of nm-sized elongated blob formed by hydrophobins encapsulating oil (represented by benzene in the simulation model), probing its stability, and wish to gain the accessibility to the size and shape of the tiny blobs. The distribution of hydrophobin coating at the oil is also investigated. For example, through the density distribution measurement of hydrophobic and hydrophilic groups of the hydrophobin, as well as oil, water with regard the center of mass of the blob, the general orientation of a blob in aqueous environment can be given. For the large blob model, which size are comparable to the minimum size determinable by the experiment, the asphericity characterized by the gyration tensor will be measured with the method introduced in several literatures [78-80]. These studies are described in detail in Chapter 2 of this dissertation.

Another targeted application is to use hydrophobins in solution processing of polymers. Solution processing is virtually required to convert many polymers, particularly those that are rodlike or semiflexible, into useful products ranging from high-strength fibers to semiconductor and opto- electronic devices. The characteristics of stiff polymers often require the use of "unfriendly" solvents and high temperatures during manufacture, which is undesirable from sustainability, environmental, cost, energy and safety perspectives. Experimental studies performed by the Russo and Reichmanis groups suggest that hydrophobins can encapsulate semiflexible polymers as stable, aqueous dispersions, and improve their performance in advanced technologies. If a class II hydrophobin solution is covered by a polymer solution, the protein encapsulates and stabilizes *cylindrical* bodies of that polymer, forming *cylindrical* structures containing polymers. Thus, through molecular simulation approaches, we could provide insights about the promotion of rodlike polymer orientation and alignment by hydrophobin at the molecular level, in order to complement experimental studies or perhaps even guide them. We first developed a Martini model for the semiflexible polymer Poly(γ-stearyl α,L-glutamate), PSLG. This polymer has low-polydispersity (Mw/Mn ~ 1.05), and can be dissolved in organic solvents such as tetrahydrofuran (THF). This model was our first step toward eventually developing a Martini model for poly(3-hexylthiophene), P3HT, a representative of the polythiophene family that is the most widely investigated and intriguing semiconducting polymeric material. Both AA and CG force fields are respectively considered to model PSLG with hydrophobin coating at the organic solvent/water interface, based on what we have learned about the synthesis, structure and molecular properties of the semiflexible polymers. For both AA and CG models, the first step will be verify that the simulations can reproduce available experimental properties such as Rg and density of the polymer and hydrophobins. Meanwhile,

the CG model is compared to AA to check whether the bonded distributions are close to each other. After the verification step, a final model of PSLG with hydrophobin coating at the organic solvent /water interface is built. As the final model has length and time scales that are significantly larger than normal model accessible to the AA, only simulations with the CG model were performed. The MD simulation can feed back to the experiment team of Russo. For example, the model can easily vary the initial alignment of the semiflexible polymer near the hydrophobin coated oil/water interface, and the effect of the initial alignment on the final orientation of the polymer. The polymer concentration and/or the organic solvent can be changed as well, acting as the impact factors for the final orientation. These studies are described in detail in Chapter 3 of this dissertation. Finally, Chapter 4 contains our conclusions and suggestions for future research directions.

Chapter 2. Modeling Hydrophobins at Gas/water and Oil/water Interfaces: Interfacial Properties Connecting to the Feasibility of Hydrophobins as Oil Dispersant

Our first project is modeling hydrophobins at gas/water and oil/water interfaces. Hydrophobins are abundant in nature and nontoxic, and it might be a good substitution to the commercial oil dispersant. However, fundamental understanding of the interfacial property of hydrophobins when they encapsulate oil or gases is required. We performed MD simulation to provide molecular level details of oil blobs, gas bubbles formed by hydrophobins. Stability of these structures can also be probed.

2.1 Method

2.1.1 Protein and oil alkane models

In these studies we considered the class I hydrophobin EAS; its structure was taken directly from the Protein Data Bank (PDB) [81] as deposited by Kwan *et al.* [30] (PDB entry ID: 2FMC). We chose this entry as our starting structure because the same research group who deposited the structure of EAS used it in their MD simulations, and further validated the structure by multiple experimental techniques including NMR [14, 36]; therefore, the PDB entry contains multiple lowest-energy states simultaneously. One of the structures was randomly selected as the starting structure for the simulation and later analysis. The PDB also gave the structure of hen egg-white lysozymes (PDB entry ID: 1AKI [82]), which was used as a control group in the interfacial tension measurement in our study. For simplicity, oil alkane was modeled as pure n-decane with the chemical formula $C_{10}H_{22}$ and a density of 730 kg/m^3. All molecular simulations were performed with the MD package GROMACS, version 4.0.7 [83]. We considered both all-atom and coarse-grained models for our simulations. At first we performed atomistic simulation for gas/water interfaces, but then found that this type of simulations for

oil/water interfaces were going to be prohibitively expensive, because the system for oil/water interface had a much larger number of molecules and thus the calculations would require unaffordable simulation times. Calculations with coarse-grained models were performed for oil/water interfaces in an attempt to overcome these limitations. But coarse-grained models were also applied for gas/water interfaces and a good agreement was found with atomistic simulation.

2.1.2 Atomistic simulations

Optimized Potentials for Liquid Simulations (OPLS) force fields [84-91], which have been used for simulations of amides since 1985 [91], are among the classic atomistic force fields that have been widely used for protein simulation [92]. Therefore, OPLS-AA force field [84] was used for our atomistic simulations, specifically to model the amino acid residues of EAS and the alkyl groups of n-decane. Water was modeled using the SPC/E force field of Berendsen *et al*. [93], while parameters of NaCl was developed by Vácha *et al.* [94] Previous studies [95-97] have proved this NaCl model was consistent with the non-polarizable SPC/E water. To simulate seawater, 3.5 wt % NaCl was added to the water. The term "water" reported in the remainder of the chapter refers to model seawater, unless explicitly stated.

Conventional MD simulation and (PMF) calculation were carried out separately. In the conventional MD simulation, an EAS molecule was initially placed in different surrounding environments under periodic boundary conditions (PBC). Though hydrophobin was highly unlike to stay in gas or pure oil, we performed these two series of MD simulation as control groups. The size of each unit box varied slightly (6 - 7 nm in x and y dimensions, while about 6 in z dimension if EAS was in single phase; about 30 nm in z dimension if EAS was in two phases), but all large enough to disallow direct interactions between periodic images. To build the starting structure for the production run, firstly, a molecule of EAS was centered in a

rectangular box that had a minimum distance of 1.5 nm between any atoms of EAS and the edge of the box. As shown in Table 1, System 1 and System 2, the box was solvated with water and oil respectively. The gas/water interface in System 4 was created by extending z-direction of System 1 by a factor of 2, according to the method of reference [36], in order that the periodic system contained half vacuum after the expansion. When constructing System 5 (water/oil/water), an extra pure water box as well as an extra pure oil box were created, both of which having the same x and y dimensions as the box of System 1. Then these three boxes were integrated into one simulation box. In System 5, the starting structure of EAS in gas was extracted from the result of simulation of EAS in bulk water, according to the method of Patriksson *et al*. [98].

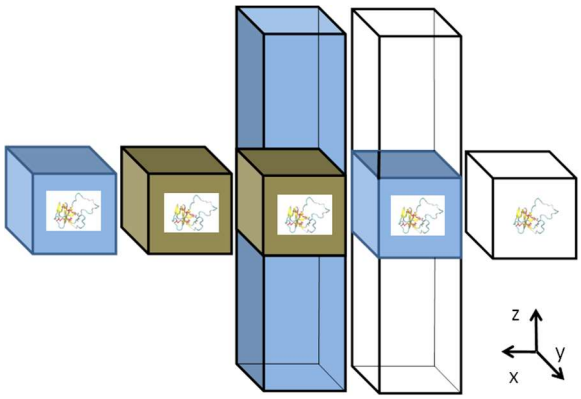


Figure 9. Schematic for the illustration of different simulation boxes. From left to right: EAS in water (blue); EAS in oil (brown); EAS in water/oil/water (blue/brown/blue); EAS in gas/water/gas (white/blue/white); EAS in gas (white). These five boxes are successively associated with System 1 – 5 in the Table 1.

Before the MD production run, each system was energy minimized with a steep descent method, and equilibrated in NVT ensemble at 298 K for 100 ps. Then molecules were further relaxed with 100 ps NPT MD simulation at 298 K and 1 bar. No barostat was applied if the system contained any gas phase. Position restraints of $1000 \cdot kJ \cdot mol^{-1}\ nm^{-2}$ were applied to every atoms of EAS except hydrogen during the equilibrium step. Temperature was maintained at the

target value using v-rescale coupling algorithm [99] with time constant $\tau_t = 0.1$, while pressure was maintained at the target value using Berendsen coupling algorithm [100] with time constant $\tau_p = 1.0$. Bonded length were constrained using the SETTLE algorithm [101] for the water and the LINCS algorithm [102] for other components. The Lennard-Jones interaction was truncated at the distance of 0.9 nm. The particle-mesh Ewald (PME) method [103] with a cutoff of 0.9 nm and a grid spacing of 0.12 nm was used for taking account of all columbic interaction except EAS in gas, where the reaction-field method was used [104, 105]. The long range dispersion corrections were applied for energy and pressure. Finally a 20 ns production run with time step of 1 fs (0.2 fs for the system having oil/water interface) was applied for each system. For the system simulating EAS in gas/water or water/oil/water, the simulation time was extended to 50 ns.

Table 1. A summary of conventional MD simulation

System	Force field	Environment	Water Molecule	Decane Molecule	Simulation Time (ns)
1	OPLS-AA	Water	8723	0	20
2	OPLS-AA	Oil	0	591	20
3	OPLS-AA	Water/oil/water	30825	875	20
4	OPLS-AA	Gas/water	8723	0	20
5	OPLS-AA	Gas	0	0	20
6	Martini	Water	7300	0	20
7	Martini	Oil	0	1301	20
8	Martini	Gas/water	15452	0	20
9	Martini	Water/oil/water	27296	1080	100
10	Martini	Gas	0	0	20

*The Martini force field uses a CG 4 to 1 mapping polarizable water to simulate water molecules.

The constrained force method [106-108] was applied for potential of mean force (PMF) calculations. In this situation, the PMF calculation can obtain the free energy profile when moving an EAS molecule along z-axis direction during multiple independent simulations with an increment of 0.1 nm. The gas/water interface plane was perpendicular to the z-axis. Based on the

free energy profile distribution, the thermodynamic preference of where EAS should stay could be found, complementing the results from conventional MD. Each of these simulations was run in NVT ensemble at 298 K for up to 2 ns; after an equilibration period of about 0.5 ns, the forces were averaged during 1.5 ns period. Other details of PMF calculations are the same as those in the paper by Liyana-Arachchi *et al*. [95, 96, 109-113]. We could only perform PMF calculations using all-atom force fields for the system having gas/water interfaces. The water/oil/water system had a much larger number of molecules and thus required prohibitively long simulation times to sample the constraint forces in an accurate manner. Calculations with coarse-grained models were performed in an attempt to overcome these limitations.

2.1.3 Coarse-grained (CG) simulations

We considered simulations with CG models to investigate systems of interest with large dimensions for longer time scales. We used the Martini force field [114-117] for EAS, water, decane, benzene and NaCl. Figure 10 illustrates a schematic for the mapping of an atomistic

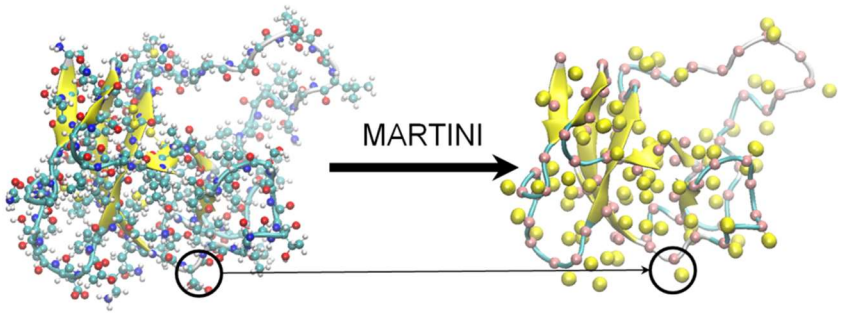


Figure 10. Schematic for the mapping of an atomistic structure of EAS (left), to a CG structure (right), conducted by the Martini force field. For clarity, the second structures of EAS represented by the NewCartoon method in VMD were included on both sides. In the coarse-grained system, the backbone of EAS was represented by pink beads and the side chains of EAS were represented by yellow beads.

structure of EAS molecule to a CG structure. Martini Version 2.2 P is used in combination with a 4-1 mapping polarizable water [118]. Because hydrophobins such as EAS are fairly rigid

molecules, the Martini force field was combined with an elastic network model named ELNEDIN [119] to conserve second and tertiary structures of EAS without sacrificing realistic dynamics of the simulation. More specifically, the backbone beads of each amino acid were constrained by springs characterized by a force constant K_{spring} of 500 $kJ \cdot mol^{-1} \cdot nm^{-2}$ and a cutoff distance Rc of 0.9 nm. These two values were in the typical range in many publications about the elastic network model [119-126], so that they were expected to have an adequate quantitative agreement with atomistic simulations.

Table 1 shows systems (6 through 10) for the CG conventional MD simulations, surrounding environment of which were comparable to the atomistic conventional MD simulations. The scenarios for preparing each box of System 5 to System 9 were the same as those used in atomistic simulations. Meanwhile, the simulation parameters for computing Lennard-Jones interaction and Coulombic interaction were the same as those published in the original Martini papers [115, 116]. Both temperature and pressure were coupled by Berendsen algorithm [100] with time constant $\tau_t = 0.3$ and pressure constant $\tau_p = 3$, respectively. The target temperature of 298 K and the target pressure of 1 bar in CG MD simulations were identical to those in atomistic simulations. Once the system was energy-minimized and well equilibrated, a production run was carried out with a time step of 20 fs and a total simulation time of 20 ns.

One of the achievements of the Martini force field is the capability to accurately reproduce free energies [116, 117, 127, 128]. This capability has been verified by de Jong et al., when their CG PMF results for amino acids residues partitioning a dioleoyl-phosphatidylcholine (DOPC) bilayer interface highly agrees with the OPLS-AA model[117]. Also, Marrink *et al.* have calculated the PMF of extracting a dipalmitoyl-phosphatidylcholine (DPPC) lipid from, and traversing a lipid through a bilayer (DPPC lipids and water) for the CG model with the Martini

force field, and then compared their results with the PMF for the atomistic DPPC model. The PMF results of both the atomistic and the CG model appear very similar [127]. We performed PMF calculation for the system having gas/water interface, which was coarse-grained by the Martini force field, and directly compared against our results obtained using all-atom models. By conducing both atomistic and CG simulations, we are able to compare thermodynamical properties obtained through atomistic and CG simulations, then we can use this "verified" CG method to probe more complicated systems that are too large to be modeled by atomistic force fields. With the same procedure as atomistic PMF calculations, one molecule of EAS was moved from bulk water to deep gas phase, crossing the gas/water interface in multiple simulations to record the constraint force needed for PMF calculation. During the equilibrium step, there is an option of short (10- 25ps) NVT MD using a 1 fs time step with position restraints on each EAS beads. This option can help further relax both the solvent molecules and the protein in the force field before the next large time step NVT equilibrium step [119]. Each production runs in these simulations was run in NVT ensemble at 298 K for 10 ns; after an equilibration period, the forces were averaged during last 5 ns period. We then conducted similar PMF calculations, but now moving the EAS hydrophobin between oil and water phases.

Interfacial tension measurement was carried out using simulations in the NVT ensemble. The same as the CG conventional MD, the Martini force field combined with the ELNEDIN network was used to model the EAS, associating with the CG polarizable water. Nevertheless, the simulation system became more complex when multiple EAS molecules were added at the oil/water interface to build a desired EAS monolayer. Therefore, only CG simulations were applied to the interfacial tension measurement. After a 20 ns production run, the surface tension γ could be calculated by:

$$\gamma = \frac{1}{2} L(P_N - P_T) \quad (1)$$

Here L is the box dimension (L=Lz in our starting structure), P_N and P_T are the normal and tangential components of the pressure with respect to the planar interface respectively. We measured a series of surface tension at different interfaces, and compared to the experimental results when available [129-131].

2.2 Results and discussion

2.2.1 PMF of EAS in gas/water and oil/water interfaces

Figure 11 displays the PMF results when moving one molecule of EAS from bulk water (where the value of PMF was arbitrarily set to 0 kJ·mol^{-1}) to gas phase. During this movement, the EAS was moved across the gas/water interface, then into the gas phase. The free energy minimum obtained from atomistic simulation is -28.9 kJ·mol^{-1}, while the minimum obtained from CG simulation is -25.0 kJ·mol^{-1}. These two values, along with the shape of the curves are close to one another, indicating that the PMF results from the Martini force field are analogous to those from OPLS-AA force field with an acceptable difference.

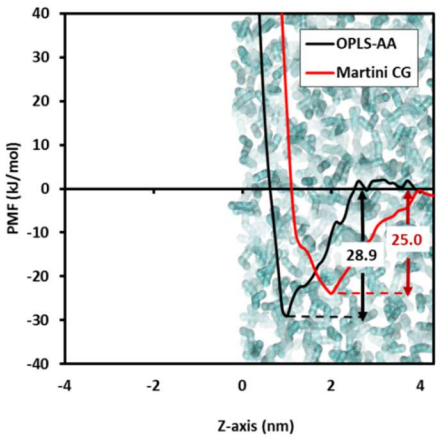


Figure 11. Potential of mean force (PMF) results. The EAS molecule was moved from the bulk water to the gas phase. Each PMF was arbitrarily set to 0 kJ·mol^{-1} in the bulk water phase and to 0 nm at the gas/water interface. Gas/water interface was defined as the point where the water density reaches 500 kg/m^3. In the background, cyan licorices represented water molecules. (a) Overall PMF (b) Zooming in the range of the PMF for clarity.

Similarly, Figure 12 displays the PMF associated with the movement of one molecule of EAS from bulk water to oil/water interface, and then back into the oil phase. Atomistic simulations were not carried out for this system as the computational cost would be very expensive. The PMF minimum is -23.6 kJ·mol^{-1} (displayed in Figure 12), which indicates EAS prefers to stay at the oil/water interface than in bulk water. In addition, once the EAS molecule is dragged out of the water phase, the PMF profile shows a sharp increase. Combining the sharp increase with the large difference between the minimum and PMF value in bulk oil, we can state that EAS is unwilling to stay in the oil phase. When EAS is at the oil/water interface, the free energy minimum of -23.6 kJ·mol^{-1} comparable to what was obtained for the case of EAS at the gas/water interface (-25.0 kJ·mol^{-1}). From Figure 12, there is a adsorption/desorption hysteresis, indicating that the true equilibrium PMF should be bracketed between the two adsorption/desorption curves. More complex simulations would be needed to sample the true equilibrium PMF, but those simulations would be expensive [132].

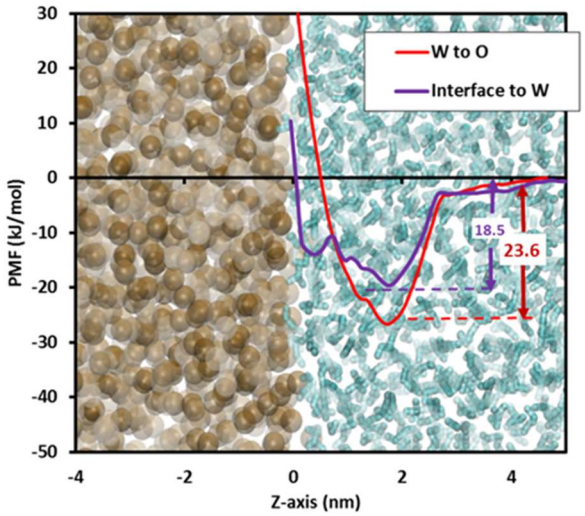


Figure 12. PMF of EAS in a bare oil/water interface. Each PMF was arbitrarily set to 0 kJ•mol^{-1} in the bulk water phase and to 0 nm at the oil/water interface. Oil/water interface was defined as the middle of two points where the water and oil density reaches 500 kg/m^3 and 500 kg/m^3 respectively. In the background, cyan licorices represented water molecules, while brown beads represented decane molecules.

Recent experiments by Reichert and Walker [133] for Tween 80 surfactant from Corexit 9500A in oil/water systems, as well as simulations by Larson research group [132] for the same system, suggest that the adsorption of this surfactant at the oil/water interface is strong and irreversible. Motivated by these results, we performed an additional PMF test with Martini force field to determine whether the adsorption of EAS to the oil/water interface is strong and irreversible. To build the simulation system, 9 molecules of EAS were coated on the oil/water interface of System 9 (Table 1) in order to simulate an EAS monolayer, fully covering the whole oil/water interface. An additional EAS was pulled from the bulk water into the oil/water interface, measuring one PMF curve related to the adsorption; then the EAS molecule

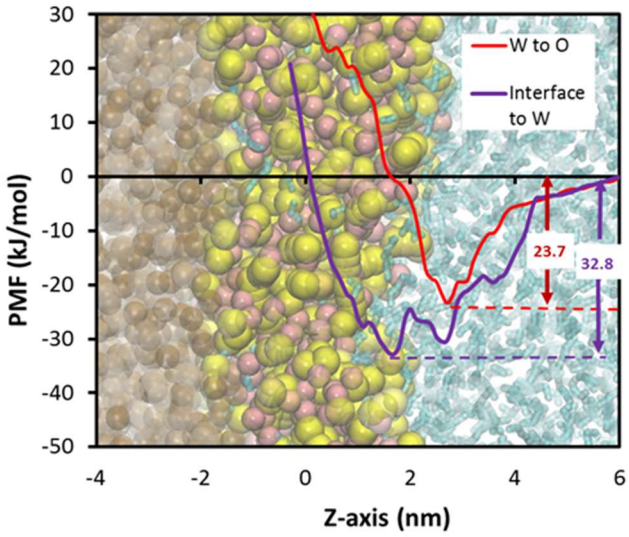


Figure 13. PMF of EAS in oil/water interface with EAS monolayer. Each PMF was arbitrarily set to 0 kJ•mol^{-1} in the bulk water phase and to 0 nm at the highest density point of EAS monolayer. In the background, cyan licorices represented water molecules, while brown beads represented decane molecules. The molecule of EAS was represented by small pink and yellow beads.

was pulled back from the oil/water interface to the bulk water, measuring another PMF curve to the desorption. As seen in Figure 13, when the interface is coated with the EAS monolayer, the PMF curve goes down until reaches at the outside of the monolayer. Then a minimum of -32.8 kJ·mol^{-1} was observed, which is somewhat lower than the case of bare interface (-23.7 kJ•mol^{-1}).

This result indicates that once the monolayer of EAS formed, more energy would be needed to remove an EAS molecule from the interface coated with a monolayer of these hydrophobins, as compared to removing an EAS molecule from a bare interface oil/water interface. The two PMF curves related to the adsorption and desorption exhibit significant differences, suggesting the presence of hysteresis upon adsorption/desorption of EAS from a coated oil/water interface. Overall, these results suggest that the binding of EAS into this interface is irreversible, in analogy to what was found experimentally by Reichert and Walker for Tween 80 at the oil/water interface [133].

2.2.2 Properties of EAS at gas/water and oil/water interfaces

Radius of gyration (Rg) of EAS in the oil/water system as a function of time is displayed in Figure 14. After EAS was bound at the oil/water interface, both atomistic and CG simulations give similar results for Rg that are close to the reported experimental Rg value of 1.25 nm [30, 134]. It is worth mentioning that in the atomistic simulation, after around 12ns (pointed by the arrow in Figure 14), the variations in the Rg of EAS become less pronounced, with the average Rg close to the experimental value. Notably, the time of 12 ns is when the EAS began to attach to the interface, suggesting that adsorption of EAS into the oil/water interface tends to stabilize the value of Rg.

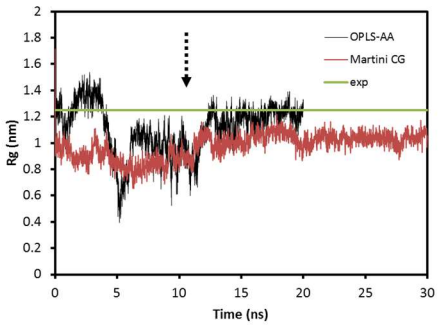


Figure 14. Radius of gyration of EAS near the oil/water interface as a function of time. Black dashed arrow pointed the time when EAS began to bind at the oil/water interface during the atomistic simulation.

Then, we introduced the measurement of Rg of the molecule of EAS at its quasi steady state, placing the molecule in different surrounding environments. During the conventional MD simulation, Rg of EAS results were obtained by averaging the Rg values of the last 15 ns. As listed in Table 2, all measured Rg from each simulation run are close to the reported experiment value of 1.25 nm. Moreover, Rg of EAS from atomistic simulation are consistently larger than from CG simulations. In addition, the Rg of EAS decreases as the hydrophobicity of the surrounding environment increases, which is expected as EAS's patches are majorly hydrophilic (Figure 15). As a consequent, when pulling the EAS molecule from bulk water to the oil/water interface, since the hydrophobicity of the environment increases, Rg of EAS decreases. This tendency could be directly reading from the Table 2: i.e., Rg of EAS in bulk water from atomistic simulation is 1.31 nm, while in gas/water interface and oil/water are 1.21 nm and 1.11 nm respectively. Furthermore, Rg is an excellent indicator of protein structure compactness [135]. Generally, α helix protein has the highest Rg and the lowest compactness. By comparison, β protein has the second highest Rg and the second lowest compactness (EAS is considered a β protein). According to the secondary structure change of EAS from bulk water to the oil/water interface, more residues are treated as β structure while no α helix structure generated. Thus, the structure of EAS at the interface is more compact than in the bulk water, agreeing well as the decreasing of Rg observed here.

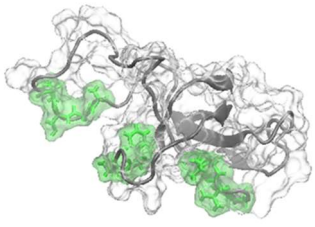


Figure 15. A depiction of the EAS in water showing the hydrophobic (green) and hydrophilic (gray) parts. For clarity, the second structures of EAS represented by the NewCartoon method in VMD were included.

Table 2. Radius of gyration of single EAS molecules

EAS in/at	Radius of gyration (nm)	
	simulation	Experiment (calculated)
Gas/water interface		1.25[a]
Gas/water interface (AA)	1.11	
Gas/water interface (CG)	1.03	
Oil/water interface (AA)	1.21	
Oil/water interface (CG)	1.04	
Gas (CG)	0.89	
Water (CG)	1.11	
Gas (AA)	1.16	
Water (AA)	1.31	
Oil (CG)	0.88	

* Experimental data are taken from Ref. [30] and Ref. [134]. Experiment used fresh water.

Furthermore, during the MD simulation, the changes in the secondary structure of EAS were calculated by the Define Secondary Structure of Proteins (DSSP) program [34, 35]. These calculations were only done for the all-atom simulations as the Martini force field is not accurate for determination of secondary structure [128]. In Figure 16 we show the secondary structure as a function of simulation time, when the EAS molecule is near the oil/water and gas/water interfaces. We also show similar results when EAS is in water, as well as results for the (hypothetical) cases of EAS being deep into the gas or oil phases. The residues forming β-structures are represented as red stripes in this figure. Though there is only one molecule of EAS placed in the simulation box, the concentration of EAS in our simulation is still higher than its possible surface saturation concentration (SSC) of 0.2-0.3 μM [58]. When the EAS structure was determined by NMR, there were four-stranded β-barrel (K15–C19, V46–V47, S57–K62, and N79–V81) and two-stranded antiparallel β-sheet (T2–T3 and Q53–C54). As the MD simulation proceeded, EAS molecule approached the interface and finally bonded at it. During this process, more β-structures could be visually found in Figure 16. The widening of β-barrels which were original Q53–C54, S57–K62, and N79–V81 contribute to the β-structures increases.

Correspondingly, different experiment techniques including circular dichroism (CD), attenuated total reflectance-Fourier transform infrared spectroscopy (ATR-FTIR) and fiber diffraction all point toward increases in β-structure upon formation of EAS rodlets at the gas/water interface [30]. In contrast, at the water/solid interface, more α-helix structure are expected to be observed from experiments[3, 20].

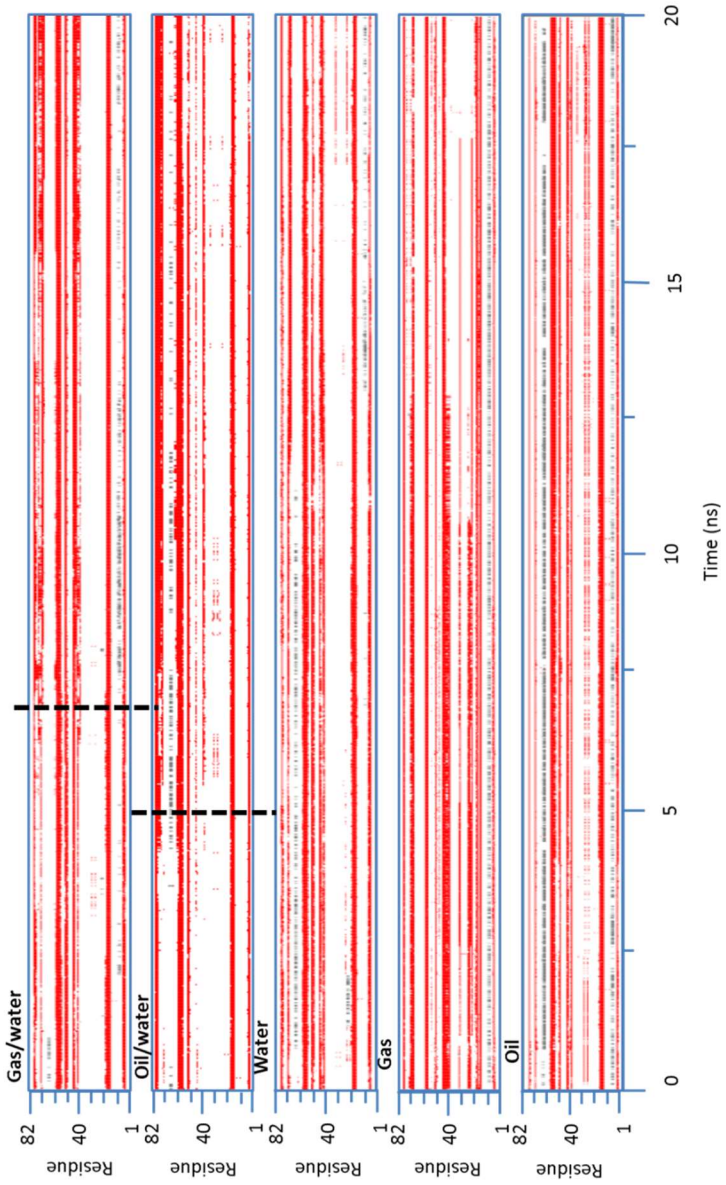


Figure 16. Secondary structure for EAS in different environments. Red represents β-structures and black represents helical structures. Black dash lines emphasized the time when β-structure increased.

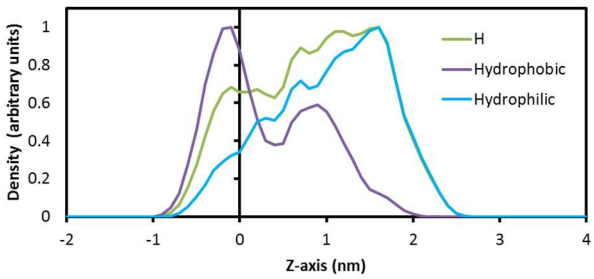
Our simulation agrees with experimental observations of the β-structure increase, indicating the analogous structures between EAS at gas/ water and oil/water interfaces. Plus, we cautiously state there is correlation between the experimental observations and our simulation on nanosecond scale about the change of β-structures, because before us some studies about the secondary structure change of small protein with MD simulation had already been carried out on this time scale [63, 98, 136], one of which used a well-studied class I hydrophobin SC3 [63].

Interestingly, when examining two control groups of EAS in pure oil and gas, some unusual structure exists. For instance, α-helix structure of EAS was observed when put in the pure oil. In another case of EAS in gas, the β-sheet of T2–T3 was not recorded by the DSSP, and some β-structures are overlapped. Since lack of data from real experiment, results of these two controls groups need to be treated with caution when considering them as structure prediction. On the whole, self-assembly of hydrophobin is accompanied conformational changes [3]. Therefore, the simulated change of EAS β-structure, which agrees with the experimental observation, would be a good indicator of the self-assembly, which may facilitate the dispersion process.

By better understanding the structure-function relations in hydrophobins, the application potential as oil dispersant could be clearer studied. When EAS stays firmly at the oil/water interface, the density profiles derived from both atomistic and CG simulations, as illustrated in Figure 17, demonstrated that EAS preferred to stay at the waterside of the interface. Moreover, as expected, the hydrophobic patches statistically determined by the HotPatch Web Gateway [137] are closer to the water side of the oil/water interface than the whole EAS molecule. Other measured parameters that will be discussed in this section are solvent-accessible surface area (SASA) and interfacial tension. Through the discussion of EAS's structural properties, especially

the surface properties at oil/water or gas/water interfaces, a better understand of its encapsulation capability of oil and gas could be achieved.

(a)



(b)

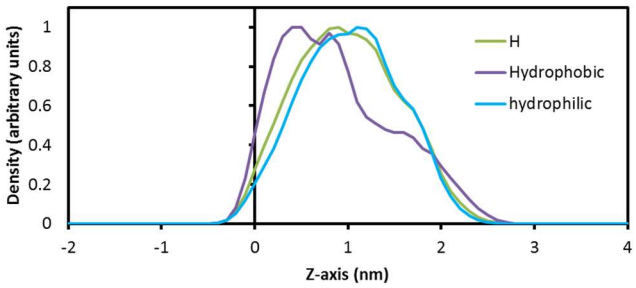


Figure 17. Density profiles for the EAS molecule in oil/water systems. The profiles are for the whole EAS molecule (green), the hydrophobic groups of the EAS molecule (purple), and the hydrophilic groups of the EAS molecule (light blue). (a) Atomistic simulation. (b) Coarse-grained simulation. The density profile of each species is normalized by dividing by the maximum value of their respective local density in the simulation box. The z coordinate was set to 0 at the middle of two points where the water and oil density reaches 500 kg/m^3 and 500 kg/m^3 respectively.

Next, from the conventional MD simulation, SASA, which is the surface area of a biomolecule that is accessible to a solvent [138], is obtained to investigate its binding style at interfaces. A well-known empirical equation [139] to predict SASA of an oligomeric protein is:

$$As \quad 5.3M^{0.760} \tag{2}$$

Here M is the molecular weight of the protein molecule. The predicted value for EAS is 49.9 nm^2 obtained from the equation above. Comparatively, the measurement of SASA is accomplished directly in the GROMACS package, applying the Shrake-Rupley algorithm [140]. The measurement results of SASA of EAS at the oil/water interface are 69-70 nm^2. The

deviation of observed SASA from its predicted value could be a useful indicator of to differentiate structural and flexible protein-protein binding, giving by the relative solvent accessible surface area [141]:

$$A_{rel} = As^{observation}/As^{prediction} \quad (3)$$

If $A_{rel} > 1.2$, in most case the binding is flexible, which means the protein is intrinsically disordered and highly flexible. We calculated A_{rel} and found a value of about 1.4, which indicates that the adsorption behavior of EAS at oil/water interface is somewhat similar to the flexible protein-protein binding. This "flexible binding" could not be directly observed from the snapshots, and it is independent from the dynamic changes of Rg indicating a more stable size of EAS upon adsorption. However, the binding type indicates there are disordered loops in the molecule EAS, for the disordered loops contribute to the high SASA and the flexibility [142]. According to Kwan *et al*'s experimental study, a portion of disordered loops are unnecessary to the hydrophobin's function [36]. Our work points the same view from the simulation side. Thus, to reduce the weight and enhance the efficiency, suitable protein mutant should be carried out first before the hydrophobin manufacture.

The last structural property of EAS presented in this section is the interfacial tension. Through the surface tension measurement, if the system interfacial tension was reduced by the addition of hydrophobin EAS, then the role of EAS as a bio-surfactant can be assured by its surface property. As mentioned in the Section 2.1, the interfacial measurement was independent from the conventional MD, because it requires the application of the NVT ensemble instead of the NPT ensemble. Moreover, the interfacial tension was only determined for our simulations with the Martini CG model. Table 3 lists interfacial tensions obtained from simulations and experiments. In order to verify the Martini force field, we first run simulations for gas/water, oil/gas and oi/water systems when no hydrophobins are present. The simulated interfacial tension

for gas/water systems is significantly different from experiments, which is a well-known issue of the water models (non-polarizable and polarizable) used in the Martini force field [115, 118]. On the contrary, we found that the interfacial tension from the Martini simulation for the bare oil/gas interface agrees well with the experimental results. For the oil/water interface, when the polarizable water model was applied in our study, surprisingly our simulated interfacial tension only had about 20% deviation from the experimental results.

Table 3. Interfacial tension of gas/decane/water interfaces at room temperature.

System (all CG)	Interfacial tension (dyn/cm)	
	simulation	experiments
Gas/water	30.5[a]	72.01[b]
Decane/Gas	23.45	23.83[c]
Decane/water	41.92	51.98[d]

*All carried out at 298 K unless denoted. For specificity, the term "decane" was used.
*Experiment used fresh water.
*a: from Ref. [118]; b: from Ref. [129]; c: from Ref. [130] at 293K; d: from Ref. [131]

By varying concentration of EAS molecules at the oil/water interface, as illustrated in Figure 18, we created a series of simulation boxes with the same dimension. Beginning with one

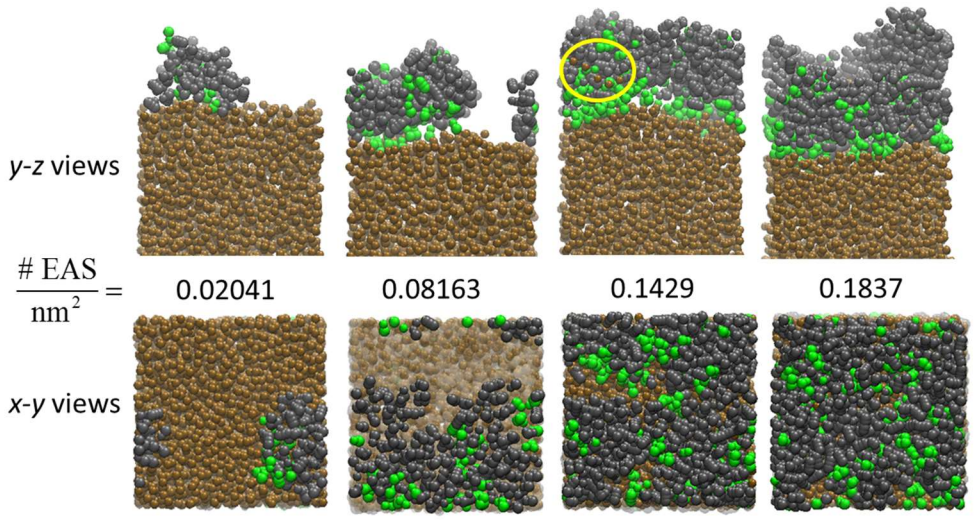


Figure 18. Conventional MD simulations by varying EAS amounts at oil/water interface. Brown beads represented decane molecules. The molecule of EAS was represented by black (hydrophilic) and green (hydrophobic) beads. For clarity, water is not shown.

EAS molecule at the interface in simulation box, more and more EAS was added to the box until the interface was fully covered by the EAS molecules. The amount of EAS varies from 1 to 9 in different simulations, changing the number of EAS·nm^{-2}. Most parts of EAS are on the water side of the interface. The hydrophobic patches in green are closer to the interfaces than the hydrophilic patches. From the snapshot we can find, as we increase the hydrophobin numbers at the interface, some oil molecules are captured within the EAS layer.

The density profile result confirmed the snapshot observations in Figure 19. As the hydrophobin coverage increases, density profiles of oil, seawater become less sharp, implying that some water and oil molecules are penetrating into EAS layer. Hydrophobin density profiles become significantly wider as well.

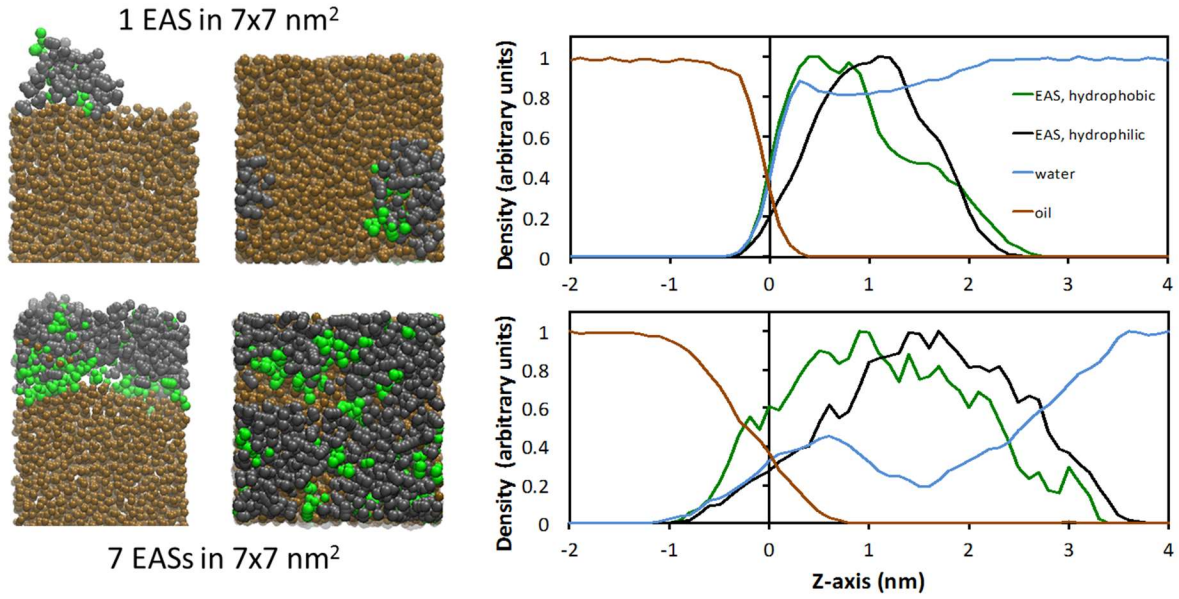


Figure 19. Density profiles change as EAS amounts varying at oil/water interface. Brown beads represented decane molecules. The molecule of EAS was represented by black (hydrophilic) and green (hydrophobic) beads. For clarity, water is not shown. The profiles are for the hydrophobic groups of the EAS molecule (green), and the hydrophilic groups of the EAS molecule (black), oil (brown) and water (blue). The density profile of each species is normalized by dividing by the maximum value of their respective local density in the simulation box. The z coordinate was set to 0 at the highest density point of EAS monolayer.

We measured the size of hydrophobins, the average size of HPs do not change as we vary hydrophobin amounts at oil/water interface (Figure 20).

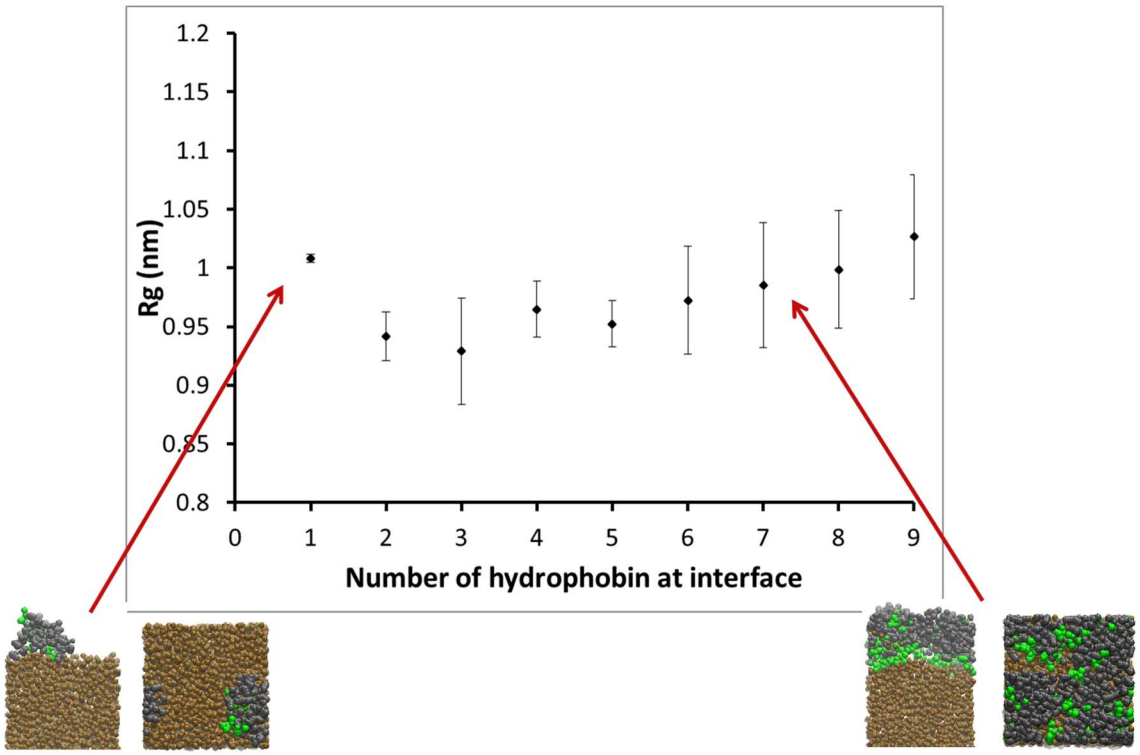


Figure 20. Size of EAS change as a function of EAS amounts at oil/seawater interface. The molecule of EAS was represented by black (hydrophilic) and green (hydrophobic) beads. For clarity, water is not shown.

We then measured the interfacial tension at oil/water interface with different EAS coverages. Increases in the surface concentration of EAS molecules at the oil/water interface leads to reductions in the interfacial tension of up to one third of the measured value for bare oil/water interfaces. Reductions in the interfacial tension are also observed when lysozymes fully coat this interface, but these reductions are not as significantly as those observed when EAS is present at these interfaces. Though these interfacial tension measurements are not accurate as they have deviations of about 20% from experimental values, their trends have qualitative agreement with contact angle measurements of interfacial tensions in similar environments [36].

These results indicate that hydrophobins can reduce the surface tension at the oil/water interface, which implies a good potential of hydrophobin as oil dispersant.

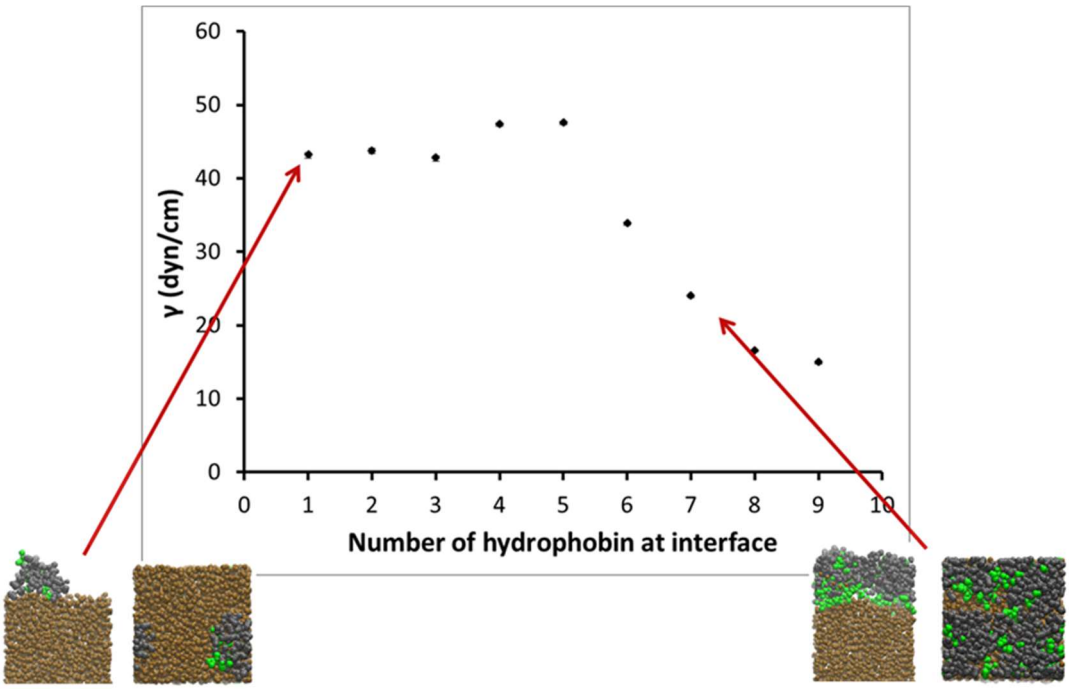
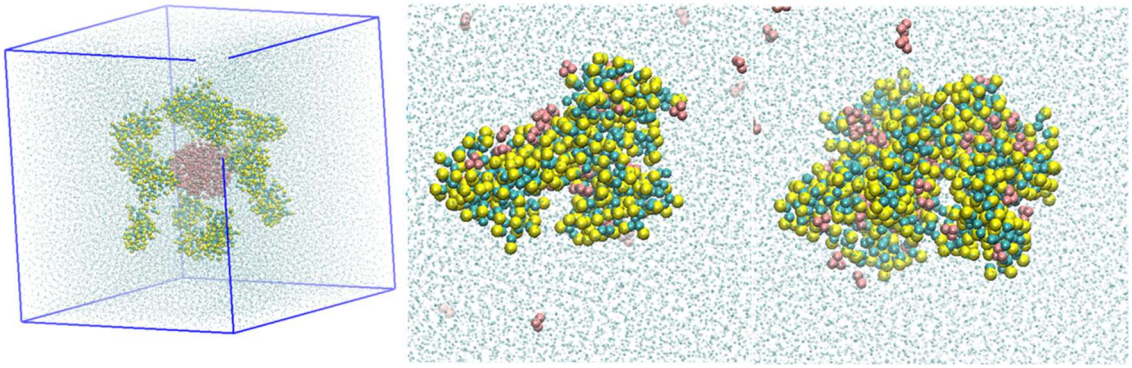


Figure 21. Surface tension decreases as we increase EAS amounts at oil/seawater interface. The molecule of EAS was represented by black (hydrophilic) and green (hydrophobic) beads. For clarity, water is not shown.

2.2.3 Characterizing nm-sized oil 'blob' properties from MD simulations

We have performed MD simulations using CG models to gain insight into the stability of the 'blobs' formed by the assembly of hydrophobin around small oil droplets. The stability of the bubbles has been measured in the simulations by monitoring root mean squared deviations (RMSDs) from the initial sizes as a function of effective simulation time. The characteristic dimensions of the bubble were monitored as a function of the effective simulation time. For system involving hydrophobin blobs, hydrophobins were initially placed around a pre-equilibrated benzene droplet with either sphere or cylinder shape in bulk water. During the production runs, the hydrophobins were expected to aggregate around the benzene droplet, forming a relatively stable structure in the water environment.

We created blobs' of hydrophobin and benzene in bulk water, following the principle of going from simple to complex. This can also probe whether hydrophobin blobs encapsulating benzene have different stabilities based on their shape and size. Table 4 shows various models we have built beginning with 0 benzene and 4 hydrophobins. As more and more benzene and hydrophobins were applied for creating larger blob model, finally a blob large enough to reach the smallest size determined by the experiment was built. We then created several independent models of large blob by varying the concentration of benzene and hydrophobins. A blob with double-layers of hydrophobin was created as well. Figure 22 displays two examples of simulation results of structures formed by hydrophobins around benzene nm-sized droplets. Blob size was stable over the course of simulation.



(a) (b)

Figure 22. The MD simulation of hydrophobins around benzene in bulk water. The molecule of EAS was represented by cyan and yellow beads, and the molecule of benzene was represented by pink beads. Even smaller dots denote water.

Besides the visual observations, there are other measurements that could help to elucidate the properties of these structures. For instance, measurements of the moment of inertia of the structures over the simulation time suggest that these 'blobs' have an ellipsoidal shape, consistent with the experiment observation of *cylindrical* blobs. Table 4 records all of the blob

models we have made, the simulation time as well as the basic size and shape information we measured.

Table 4. Blobs' of hydrophobin and benzene in bulk water

Benzene	Hydrophobin	Sim. time (ns)	Avg. Rg (nm)	Moment of Inertia
0	4	500	2.4	1:2.2:2.7
0	8	500	3.6	1:2.9:3.4
48	4	1000	2.1	1:1.1:1.7
32	4	1000	2.1	1:1.2:1.5
21	4	1000	2.0	1:1.1:1.2
64	8	1000	2.3	1:2.1:2.4
82	8	1000	2.2	1:1.3:1.7
212	14	1000	4.5	1:1.9:2.3
466	24	1000	5.0	1:1.6:1.7
3334	80	2000	9.4	1:1.5:1.7
3334	109	2000	9.2	1:1.3:1.4
10169	109	2000	8.0	1:1.1:1.1
3334	218	2000	11.8	1:1.2:1.4

We finished performing production runs with larger system sizes that are comparable to real blobs observed in experiments by the Russo group. As illustrated in Figure 23, 109

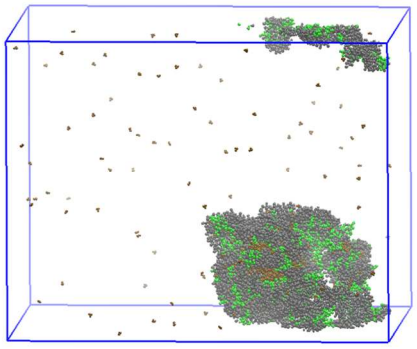


Figure 23. The MD simulation of 109 hydrophobins around 3334 benzene in bulk water (2231196 water molecules). The size of the unit box is 38 * 38 * 48 nm. The molecule of EAS was represented by black (hydrophilic) and green (hydrophobic) beads, and the molecule of benzene was represented by brown beads. For clarity, water is not shown.

hydrophobin and 3334 benzene molecules were solvated in water. There were 2231196 water molecules in the unit box with a size of 38 * 38 * 48 nm. The EAS molecules were initially placed around the benzene. During a 2000ns conventional MD simulation, a *cylindrical* blob was formed. The RMSD of the whole blob with respect to its initial structure test suggested that the blob is stable, as shown in Figure 24.

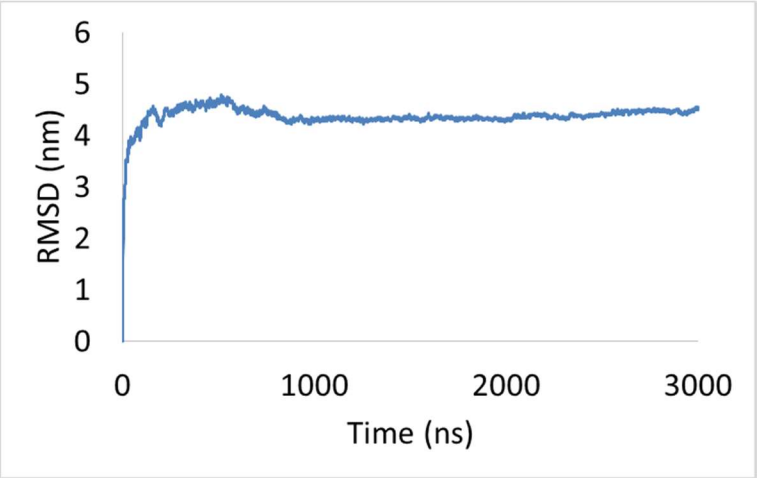


Figure 24. Root mean square deviation of the whole blob with respect to its initial structure. There are 109 hydrophobins around 3334 benzene in bulk water。

We then created several independent models of large blob by varying the concentration of benzene and hydrophobins, based on the "normal" blob introduced above (109 hydrophobins around 3334 benzene). A lower hydrophobin concentration blob (80 hydrophobins around 3334 benzene), as well as a higher benzene concentration blob (109 hydrophobins around 10169 benzene) were created respectively. A blob with double-layers of hydrophobin was created as well (218 hydrophobins around 3334 benzene). All large blobs were formed and stable short after the production run began. The base size and shape of the blob were recorded. Density of selected groups in the oil blobs as a function of their distance from blob's center of mass distance were measured, and the shape of the density distributions are in good agreement with the visual check. Figure 25 shows there are very few water molecules inside the blob, meanwhile, the

benzene molecules stayed in center. The hydrophobic and hydrophilic patches of hydrophobin determined by the HotPatch Web Gateway were battlefield verified through the distribution test, as the determined hydrophobic stayed closer to the benzene, while the hydrophilic stayed closer to the water.

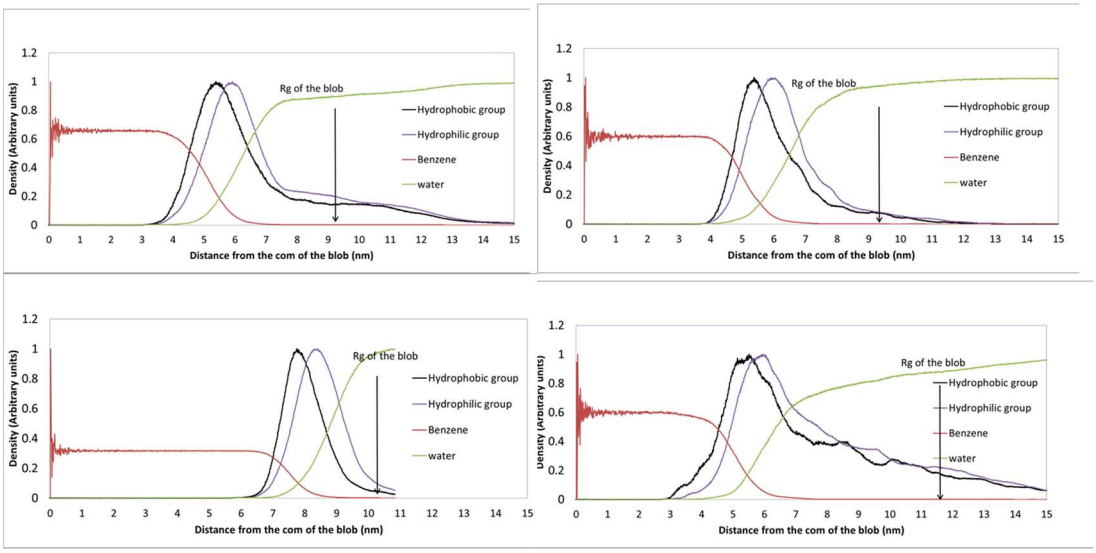


Figure 25. Density distribution for different groups of the oil blobs in terms of their distance from blob's center of mass distance. The selected groups are waters in green, hydrophilic patches in blue, hydrophobic patches in black and benzenes in red. The black arrows indicate the Rg of the blob. Left top: 109 hydrophobins around 3334 benzene; right top: 80 hydrophobins around 3334 benzene; left bottom: 109 hydrophobins around 10169 benzene; right bottom: 218 hydrophobins around 3334 benzene.

Table 5 shows the size of blob in Rg and shape in moment of inertia. For large blobs, the size and shape not only can be calculated directly by Gromacs, but also can be derived from the gyration tensor measurement [79], which is considered as a more suitable and accurate method for characterizing such complex model. Hall *et al.* [80] adopted the gyration tensor concept and applied it for their large oil aggregates measurement. This method involves calculating the principal components of the gyration tensor, S, for each target object. In the equation below, x_i, y_i, and z_i are the coordinates of each hydrophobin or benzene molecule, i, in the aggregate, x_{cm}, y_{cm}, and z_{cm} are the coordinates of the hydrophobin or benzene's center of mass, and N is the total

number of hydrophobin or benzene in the blob. The eigenvalues of the gyration tensor matrix, λ_1, λ_2 *and* λ_3 are principle components of Rg for size measurement, and can then be used to characterize asphericity for shape measurement, respectively. The asphericity has a minimum value of 0 for perfectly spherical aggregates and a maximum value of 1 for completely linear blobs. The last 500 ns out of total 2000ns simulations were used to calculate the gyration tensor matrix.

$$\boldsymbol{S} = \frac{1}{N}\begin{pmatrix} \sum_i (x_i - x_{cm})^2 & \sum_i (x_i - x_{cm})(y_i - y_{cm}) & \sum_i (x_i - x_{cm})(z_i - z_{cm}) \\ \sum_i (x_i - x_{cm})(y_i - y_{cm}) & \sum_i (y_i - y_{cm})^2 & \sum_i (y_i - y_{cm})(z_i - z_{cm}) \\ \sum_i (x_i - x_{cm})(z_i - z_{cm}) & \sum_i (y_i - y_{cm})(z_i - z_{cm}) & \sum_i (z_i - z_{cm})^2 \end{pmatrix} \quad (4)$$

$$R_g^2 = \lambda_1 + \lambda_2 + \lambda_3 \quad (5)$$

$$A = \frac{(\lambda_1 - \lambda_3)^2 + (\lambda_2 - \lambda_3)^2 + (\lambda_1 - \lambda_2)^2}{2(\lambda_1 + \lambda_2 + \lambda_3)^2} \quad (6)$$

Table 5 Radius of gyration and asphericity characterized by gyration tensor

Benzene	Hyd	Rg. (benzene)	Asphericity (benzene)	Rg. (hydrophobin)	Asphericity (hydrophobin)
3334	80	5.59	0.0047	7.53	0.2
3334	109	6.60	0.080	12.32	0.33
10169	109	7.12	0.034	9.07	0.086
3334	218	5.87	0.022	10.71	0.14

From the results shown in the above table, the sizes of the simulated blobs were all comparable to the experiment. The overall shape of each blob, which is essentially the same as the shape of the hydrophobin layer, is slightly ellipsoidal except for the blob having large amount of benzene. The benzene molecules in each blob adopted a more spherical shape than the overall blob, whereas the hydrophobin layer in each blob made the blob to adopt a slightly ellipsoidal shape. The results implied the ratio of hydrophobin and the organic solvent encapsulated by them played important role on the overall shape of the blob. When there are double layers of hydrophobin coated on the benzene, the shape of blob became more spherical.

2.3 Research progress conclusion

We have investigated the properties of EAS, a class I hydrophobin, near gas/water and oil/water interfaces. According to the PMF calculation results, EAS molecule is likely to stay at the interfaces as indicated by the presence of deep free energy minima. Conventional MD simulations also confirm that EAS prefers to stay at the water side of the interfaces, and once it is adsorbed EAS did not leave the interfaces. This is consistent with the thermodynamical trends we observed from PMF calculations. This preference implies the good capability of hydrophobin to stabilize the hydrophobic/hydrophilic interface as a bio-surfactant. Our PMF calculations also suggest that the adsorption behavior of EAS at the interface is strong and irreversible.

After EAS binds to the interfaces, the value of Rg was close to the experimental value and experience small variations. The overall structure of EAS became more stable at the interface than in the bulk water. In addition, the secondary structure change of EAS was consistent with the experimental observations, indirectly proving the self-assembling process occurs at the oil/interface, which is a major impact factor for dispersing oil in water.

The decrease of Rg when moving the hydrophobin from bulk water to interface agrees with the increase of secondary structure of EAS, in analogy to the self-assembling. Last, the interfacial tension of the oil/water interface had been reduced significantly by adding EAS, suggesting that hydrophobin is a powerful surfactant. Those capabilities promised hydrophobins have a good potential as oil dispersant.

Besides the hydrophobin-coated interface simulations, we conducted MD simulations with Martini CG FF to gain insight into the stability of nm-sized ‘blobs’ formed by the assembly of hydrophobin around oil. The class I hydrophobin EAS molecules were initially placed around a pre-equilibrated benzene blob with *cylindrical* shape in bulk water, then the production run

lasted for 500 -2000ns. We then increased the blob size, and finally built a large blob which size was comparable to the smallest blob from experiment. The hydrophobin molecules were successfully assembled at the benzene blob surface, and the blob size was stable over the course of simulation. Most blobs had a slightly elongated shape, as determined through moment of inertia and asphericity measurements.

Chapter 3. Molecular Dynamics Simulations of the Interface between Semiflexible Polymers and Hydrophobins Encapsulating Them

The results presented in the previous chapter suggest that CG models in MD simulations might capture the essential physics behind the striking surface activity of hydrophobins, hence could be used to explore the interfacial properties of systems of semiflexible polymers, organic solvents, water and hydrophobins.

Solution processing is virtually required to convert many polymers, particularly those that are rodlike or semiflexible, into useful products ranging from high-strength fibers to semiconductor and opto-electronic devices. The characteristics of stiff polymers often require the use of "unfriendly" solvents and high temperatures during manufacture, which is undesirable from sustainability, environmental, cost, energy and safety perspectives. Experimental results from the Russo and Reichmanis groups at Georgia Tech have demonstrated that hydrophobins can encapsulate semiflexible polymers as stable, aqueous dispersions, and improve their performance in advanced technologies. The focus of semiflexible polymers is Poly(γ-stearyl α,L-glutamate), PSLG, of which low-polydispersity (Mw/Mn ~ 1.05) can be easily obtained in convenient solvents such as tetrahydrofuran (THF). Although trails the helical polypeptide PSLG in terms of rigidity, monodispersity and solubility, poly(3-hexylthiophene), P3HT, as a representative of the polythiophene family, is also a good option because it is the most widely investigated and intriguing semiconducting material. P3HT has been widely investigated for device applications due to its hole transport properties, solubility in a range of organic solvents and good film-forming characteristics [143]. Here, molecular dynamics simulations of the interface involving semiflexible polymers in organic solvents, and hydrophobins in water, were performed. The objectives of the MD simulations are to provide insights on the interfacial properties of these systems, with the long-term aim of elucidating the molecular mechanisms that

drive the encapsulation, alignment and packing of semiflexible polymers by hydrophobins, and ultimately suggesting ways to optimize the capture and alignment of the polymers in these systems. As all-atom simulations of these systems are computationally very expensive, we developed Martini models for PSLG and the organic solvent THF; Martini models for hydrophobins were described in the previous chapter.

3.1 Introduction to PSLG

Poly(γ-stearyl α,L-glutamate), PSLG is similar to its close cousin Poly(γ-benzyl α,L-glutamate), PBLG, in which both of them have R-helical backbone. However, PSLG's properties include electrical neutrality, rigidity, polydispersity, and good solubility in several solvents [144].

This introduction focuses on the formation, structure, rod-like properties, and solution properties of PSLG. Different technology and methods are used to analyze the polymer's properties. For example, differential scanning calorimetry (DSC) and temperature-ramped fluorescence photobleaching recovery (TRFPR) are used to find its rod-like properties. Static light scattering (SLS) and dynamic light scattering (DLS) are used to find its solution properties. Even with these analyses, many PSLG's properties are still unknown and require more in-depth research.

3.1.1 Synthesis of PSLG

The formation of the complex protein structure PSLG is a long process. However, similar to other complex polymer, it all is made with small and simple molecules. First, the L-glutamate acid reacts with stearyl alcohol (octadecanol), producing γ-stearyl α,L-glutamate amino acid (SLGAA). The amino acid goes through selective esterification of the γ -carboxylic acid position by reacting with stearyl alcohol [145]. As a result, SLGAA is the start point of PSLG.

After SLGAA has been created, the next step closes the ring structure of two alcohol functional groups to form N-carboxyanhydride (NCA), which gives rise SLGNCA, which is the monomer of PSLG. Initially, the phosgenation method was used for this synthesis. However, due to the high toxicity of phosgene gas, Daly and Poché came up with a substitute compound, bis(trichloromethyl) carbonate (triphosgene), made by exhausting the chlorine in dimethyl carbonate . In order to ensure that the safer approach will lead to the same compound, C and H NMRs were used. Indeed, both NMR results showed the same structure for the organic compound. SLGNCA has properties that are highly sensitive to heat and moisture, which is similar to other NCA derivatives. However, the key difference between SLGNCA and NCA derivatives is that SLGNCA can form a white crystal in the presence of benzene, while the derivatives can only form liquids [145]. At this point, the tail of the SLGNCA (PSLG) is well developed in comparison to the ring structure, which requires phosgenation due to its functional group [145]. Lastly, polypeptide bonds are needed for the monomers to be able to combine with each other and form a protein-like structure.

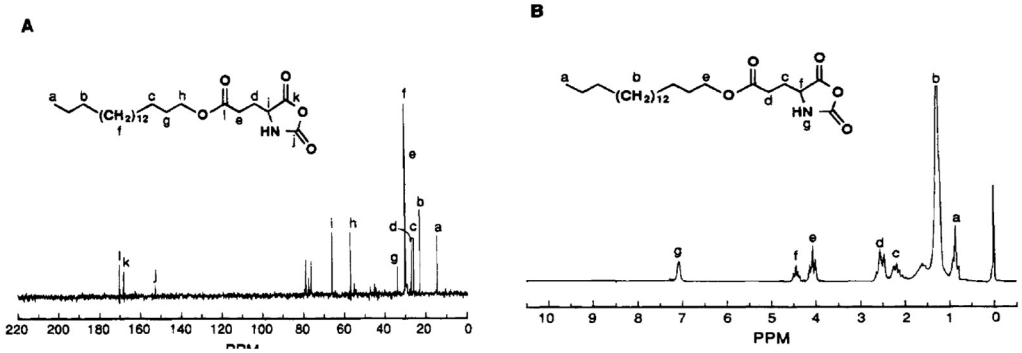


Figure 26. NMR spectra of SLGNCA. $CDCl_3$ was used as solvent at TMS standard. (A) 25-MHz ^{13}C (B) 100-MHz ^{1}H. Reprinted with permission from Ref [145]. Copyright 1995 American Chemical Society.

After the PSLG monomer is created, the next step links up all of the monomers into a single long chain. The initiation step is often used by primary and tertiary amines or sodium

methoxide. Sodium methoxide was used due to its strong base properties, which can lead to high molecular weight polymers. However, the popular choice would be primary amines since the product produces similar and predictable weight results. The recommended monomer/initiator ratio should be below 100. If it is over 100, the polymerization process becomes significantly slow with unequal weight distribution [145].

The mechanism of replicating the polymer, PSLG, begins with the initiator, a primary amine, covalently bonding to a monomer. This causes the acidic hydrogen on the nitrogen in the monomer's ring to abstract. The monomer becomes active and attacks the nucleophile site at the NCA ring's C5 position. The monomer's ring opens up. The amide and ester section of the ring loses a carbon and 2 oxygens in the form of carbon dioxide. The newly formed amino group attacks nearby monomers, beginning a chain formation of a polymer. The mechanism of this process can be found in Figure 27. [146]. Once the monomer SLGNCA has been used up, the coupling of PSLG polymer chains begins to occur. As a result, both the PSLG's molecular weight and polydiversity increase [145].

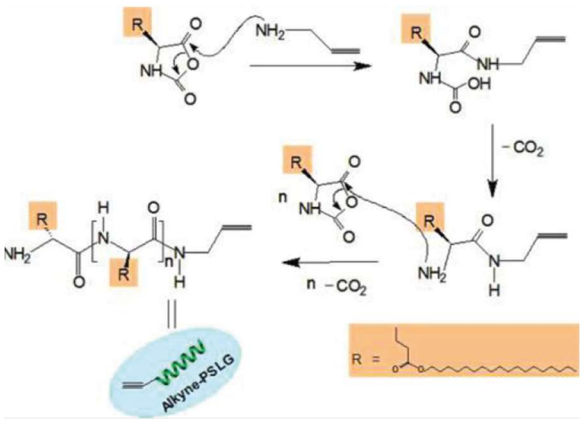


Figure 27. Polymerization of SLGNCA initiated by primary amine to form PSLG. Reprinted with permission from Ref [146]. Copyright 2010 American Chemical Society.

The factors that lead to a higher molecular weight of the polypeptide are the usage of a stronger base, higher monomer concentration, and longer "aging" time. As mentioned in the

previous paragraph, strong bases such as sodium methoxide can lead to heavier molecular weight. Methanol is also present due to the fact that sodium methoxide needs to be in methanol solution. In addition, methanol can initiate a chain, but at a much slower pace. In order to show the significant results of heavier polymer weight, the ratio of monomer to methoxide needs to be greater than 100. Moreover, a highly concentrated solution of monomer is needed as well. According to an experiment comparing a 2 day reaction to a 5 day reaction in which PSLG was created at the same initial monomer and initiator concentration, the 5 day reaction produced a polymer that was significantly heavier than the 2 day reaction [145].

3.1.2 Structure of PSLG

The primary structure of PSLG is introduced first. As mentioned above, PSLG is a polymer that rises from the monomer SLGNCA, with the only difference being the loss of one carbon and two oxygens in the polymerization process. Figure 28 is an image of a repeating unit of PSLG in its single form [144]. The secondary amide in the PSLG allows it

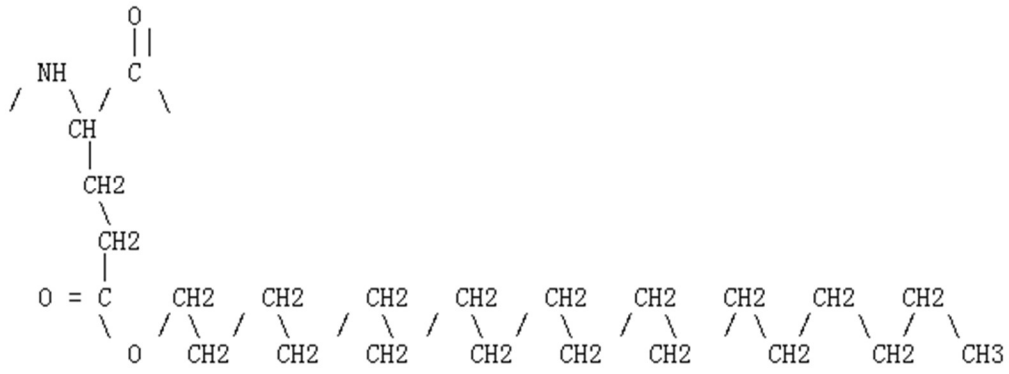


Figure 28. A single repeat unit of PSLG

to hydrogen bond with its long carbon chain since there are electronegative charges around nitrogen and oxygen. Moreover, the π-bonding in the carbonyl allows oxygen to become an

electron acceptor, while the hydrogen in the N-H bond becomes an h-bond donor. Figure 30 represents the dipolar movement of an amide functional group from a single unit of PSLG.

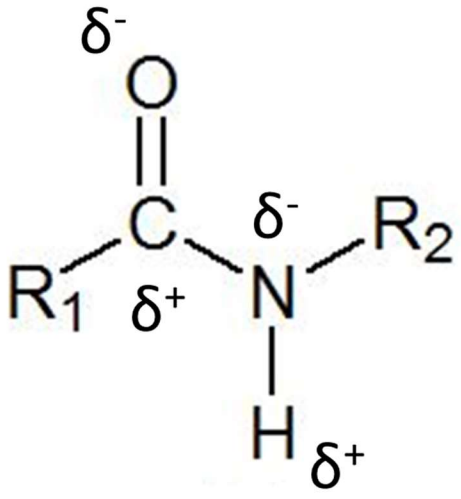


Figure 29. Dipolar nature and water interaction for amide functional group

The other functional group in the PSLG is ester. Ester is a slightly weaker electronegative functional group in comparison to amide. In addition, the long hydrocarbon chain attached to the ester weakens the electronegative properties of the functional group. Nonetheless, it still promotes polar attraction in the region.

Separating from its cousin, poly (γ-alkyl α,L-glutamate), PSLG has its long, distinct hydrocarbon chain which makes the molecule less soluble in water and more soluble in organic solvents such as toluene and THF. Moreover, the hydrocarbon chain causes PSLG to form an α-helix chain [146]. This leads us to the next discussion on the secondary structure of the polymer.

The secondary structure of PSLG focuses on the α-helix chain, which depends on environmental parameters. These environmental parameters include the solvent and operating temperatures. In either THF or chloroform solvents, PSLG tends to form a right-handed α-helix. Typically, the bare helix chain has a diameter of 5.6 Å; however, if it includes the stearyl chain,

the diameter falls between 18 and 36 Å [146]. For a full turn, the helix requires 3.6 monomers and has a spread of 5.4 Å [146]. The amide group plays a crucial role within the helical structure of PSLG and gives rise to the molecule's rod-like properties [146]; a minimum of 10 PSLG monomers can produce a stable helix structure. In the "click chemistry" method, between 18 and 19 monomer units are required to form a stable helix [146].

3.1.3 Molecular and solution properties of PSLG

Because of its α-helix backbone, PSLG forms a rod-like structure. Multiple rod-like structures can form liquid crystals. The PSLG's liquid crystal considers both lyotropic and thermotropic states since the liquid crystal can be affected by the monomer's concentration and the environmental temperature. At high temperatures, PSLG dissolves its side chain in linear alkanes, while at low temperatures, PSLG forms a gel structure. At a low concentration, PSLG goes into the isotropic phase, since it does not undergo intra-chain looping. This simplifies PSLG's formation, thus, allowing researchers to understand the concept of random coil gels. There have been two different analyzed methods conducted on PSLG: differential scanning calorimetry (DSC) and temperature-ramped fluorescence photo-bleaching recovery (TRFPR) [144].

DSC is a thermo-measurement device that can characterize phase transitions as a function of temperature and energy. The results of DSC confirms two important thermo-properties for PSLG. First, there is a significant difference between cooling and melting temperatures across all concentrations of PSLG. Second, there is no significant difference between concentrations for melting enthalpy. Figure 30 illustrates that the difference between melting and cooling temperatures falls between 13-16 ℃. In addition, the more ordered (anisotropic) liquid crystal has a 3-4 ℃ higher melting and cooling temperature than less ordered (isotropic) liquid crystal.

Figure 30 shows the calculated melting enthalpy for PSLG, falling between 30 and 42 mJ/(mg PSLG). This large gap in error shows no correlation between concentrations of PSLG and melting enthalpy [144].

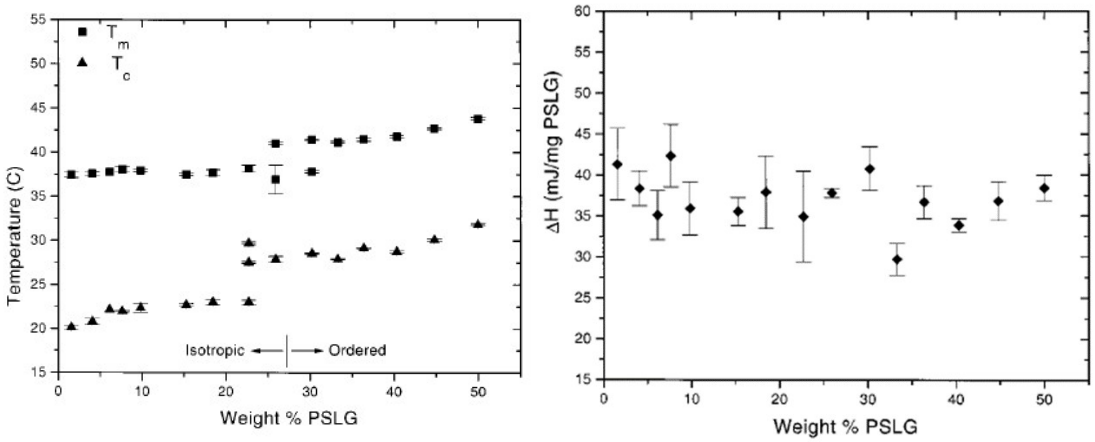


Figure 30. Various concentrations of PSLG in dodecane gels. T_m *and* T_c are melting and cooling temperatures respectively. ΔH is the melting enthalpy. Reprinted with permission from Ref [144]. Copyright 2000 American Chemical Society.

The results of temperature-ramped fluorescence photo-bleaching recovery (TRFPR) provide more accurate data on phase transition temperature compared to DSC while determining the diffusion coefficient of PSLG. TRFPR is a molecular mobility measurement device that measures the disappearance rate of fluorescent labels in the polymer with temperature variations. The fluorescent indicator weakens as the gel melts, and this fading process can capture the polymer's melting properties. The fluorescent pattern in the polymer creates alternative current (AC) and direct current (DC) signals, and the ratio between AC and DC represents the presence of the polymer. Figure 31 indicates that melting occurs at 30.9 ℃, and the diffusion coefficient is 10^{-8} cm^2/s [144]. In conclusion, both TPFPR and DSC show the phase transformation for liquid crystal PSLG; however, TPFPR can capture a more accurate temperature, making it the more popular choice for research.

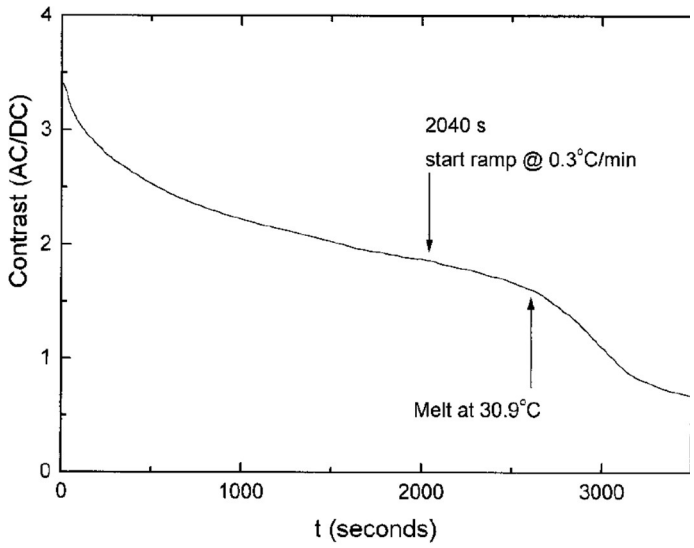


Figure 31. TFPFR trace for 10% PSLG and Dodecane gel. Reprinted with permission from Ref [144]. Copyright 2000 American Chemical Society.

The study of PSLG solution depends heavily on polydispersity, which is the weight distribution of PSLG. Two current methods to analyze polydispersity are static light scattering (SLS) and dynamic light scattering (DLS). SLS measures the average radius of gyration; this radius of gyration is linked with the molecular weight for PSLG. DLS measures the mutual diffusion coefficient, which is the ratio of thermodynamic driving force to hydrodynamic friction resistance. With this known ratio, the higher the value, the better the polymer dissolves in the solvent. However, just like SLS, these DLS values depend on the polymer weight of PSLG. Figure 32 shows the molecular weight and radius of the gyration (R_g) and radius of hydrodynamic (R_H) correlation, having a linear correlation coefficient of 0.982 and 0.999 respectively. Figure 33 shows different molecular weights concentrations that correlate with diffusion. The results show almost no correlation between concentration and diffusion, while displaying a negative correlation between molecular weight and diffusion. In conclusion, the study for solution properties for PSLG is still in progress, since only the polydispersed polymer has been analyzed, instead of the monodispersed polymer of PSLG.

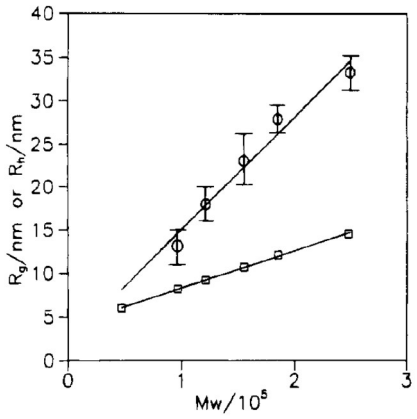


Figure 32. Radius of gyration (R_g) and radius of hydrodynamics (R_H) increase linearly with molecular weight (M_w). Reprinted with permission from Ref [145]. Copyright 1995 American Chemical Society.

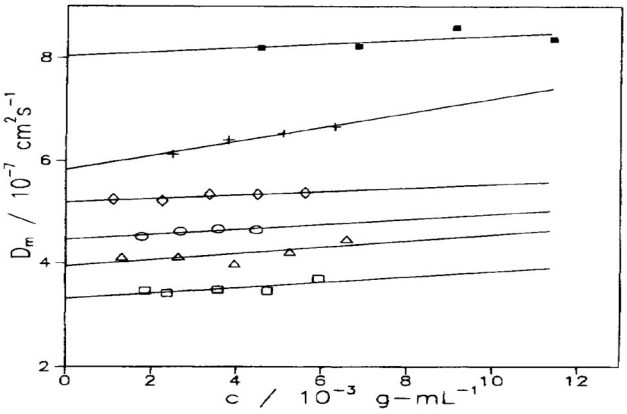


Figure 33. Concentration (c) and molecular weight depend on diffusion (D_m). Reprinted with permission from Ref [145]. Copyright 1995 American Chemical Society.

3.2 PSLG and THF models

We have systematically studied the synthesis, structure and solution properties of PSLG. The knowledge of PSLG synthesis helped us to better understand the important function of the repeat unit in PSLG. The obtained knowledge of PSLG's structure lead us to gain good insight on the functional groups of the PSLG, helped us assign the Martini beads representing those groups more clearly. Then the knowledge of its solution properties let us understand the behavior of PSLG when we change the solvent, especially THF.

The modeling process used atomistic (All-Atom, AA) and coarse-grained (CG) force field respectively. To the best of our knowledge, at the time of performing this work, there were no AA or CG models of PSLG available. EAS hydrophobin has solution structure available from the Protein Data Bank [81], and the AA model of EAS with OPLS-AA force field can directly be generated. Also, if the crystal structure is known and all amino acids forming the target protein belongs to the 20 common amino acids, the Martini force field can be used to build a CG model of EAS in Gromacs. Unfortunately, unlike the process of EAS model building, both AA and CG models of PSLG had to start with arbitrarily assigning atom (group) types to each atom or group because the structure of PSLG differs from the polypeptides made of 20 common amino acids. The initial models were then subject to some extra model performance tests to ensure the reliability of the model. To reproduce some solution properties of PSLG introduced by Poche *et al*. [145], THF was used as the solvent of PSLG. Both OPLS-AA and Martini-CG parameters of THF have been respectively applied for the MD simulations.

3.2.1 Atomistic parameterization of PSLG and THF

The AA model of PSLG was built based on what we learned about the synthesis, structure and molecular property of PSLG. Optimized Potentials for Liquid Simulations (OPLS) force fields [84-91], which have been used for simulations of amides since 1985 [91], are among the classic AA force fields that have been widely used for protein simulation [92]. Therefore, OPLS-AA force field [84], which is in the OPLS force field family, was used for our AA simulations, specifically to model the repeat units of PSLG and THF. OPLS-AA parameters of THF were developed by Jorgensen *et al.* [147] and Caleman *et al*. [148]. To reproduce experimental free energy of hydration, the THF charge is increased by 20%.

Figure 34. Atom type in a PSLG repeat unit. N represents nitrogen; H represents hydrogen; C represents carbon and O represents oxygen. The number of each atom began at the N-H bond and carbonyl and then along the long hydrocarbon chain.

The atoms of a PSLG repeat unit are shown on Figure 34. Each atom was named according to the general rule of element symbol (i.e. N = nitrogen, H = hydrogen, C = carbon

and O = oxygen) and then numbered starting at the N-H bond and carbonyl until. The numbering was then along the long alkane chain. For example, in Figure 34, the number "1" in "C1" means the first carbon atom in the whole repeat unit; the number "32" in "H32" has two meanings: while "3" means that the hydrogen is on the C3, "2" means that it is the second hydrogen of the C3. Thus, the number in C, N and O only refers to the atom position; the number in H not only refers to the atom position, but also differentiates every H on the C.

Next, since every atom in the repeat unit of PSLG already had a unique name, the atomtype for each atom could be manually defined according to the OPLS-AA regulations. Non-bonded parameters used in AA simulation of P3HT was shown in Table 6. In this table, nonbonded parameters provided by Siu *et al* [149] for the long alkane chain were implemented to the standard OPLS-AA regulations. According to the work of same authors, the standard OPLS parameters for hydrocarbons developed using short alkanes cannot reproduce the liquid properties of long alkanes in molecular dynamics simulations.

Table 6. OPLS-AA nonbonded parameters for PSLG

atom type	σ (nm)	ε (kJ/mol)	*q(e)*
N1	0.325	0.711	-0.3000
H11	0.000	0.000	0.3300
H12	0.000	0.000	0.3300
H13	0.000	0.000	0.3300
C2	0.350	0.276	0.2500
H2	0.250	0.126	0.0600
C1	0.375	0.439	0.5000
O1	0.296	0.879	-0.5000
C3	0.350	0.276	-0.1200
H31	0.250	0.126	0.0600
H32	0.250	0.126	0.0600
⋮	⋮	⋮	⋮

The torsional parameters in the form of Fourier dihedral for hydrocarbons were also updated based on the same paper, as shown in **Error! Not a valid bookmark self-reference.**. Besides optimization for the long alkane chain, a new torsional parameter in the form of improper dihedral was derived from OPLS-AA CT_2-CT-CT-C(O) for carbons beginning from beta carbon to ester carbon (C2-C3-C4-C5 in Figure 34) to keep the ester groups planar.

Table 7. New torsion parameters

Dihedral type	C_0	C_1	C_2	C_3	C_4	C_5
C2-C3-C4-C5[a]	-4.2342	7.2216	1.9079	4.8953	0.0000	0.0000
C7-C8-C9-C10[b]	0.5188	-0.2302	0.8968	-1.4913	0.0000	0.0000
C8-C9-C10-C11[b]	0.5188	-0.2302	0.8968	-1.4913	0.0000	0.0000
C9-C10-C11-C12[b]	0.5188	-0.2302	0.8968	-1.4913	0.0000	0.0000
⋮	⋮	⋮	⋮	⋮	⋮	⋮
C19-C20-C21-C22[b]	0.5188	-0.2302	0.8968	-1.4913	0.0000	0.0000
C20-C21-C22-C23[b]	0.5188	-0.2302	0.8968	-1.4913	0.0000	0.0000

Torsional potential is defined by Ryckaert-Bellemans function: $V_{rb}(ijkl) = \sum_{n=0}^{5} C_n(\cos(\varphi))^n$
The unit of C_n is kJ/mol
a was derived from CT_2-CT-CT-C(O) in the standard OPLS-AA.
b was from [149]

After all bonded and nonbonded parameters required to define the PSLG repeat unit became available, the parameters were then be saved by writing the atomtype, bonds, angle and dihedral parameters into the amino acid residue database of Gromacs. Originally the database file contains building blocks (amino acids) for proteins, but since a repeat unit of PSLG can be treated as a special amino acid, the residue database file could be used to store parameter information (i.e. bonds, charges, charge groups, and improper dihedrals) of the PSLG repeat block for the later use of pdb2gmx program. The atomtype for each atom was defined according to the standard OPLS-AA regulations and the structure of PSLG. For example, the atom C5 was defined as opls_465 with a charge of +0.51, which is an atomtype for ester carbon provided in OPLS-AA. Two newly added torsional parameters in the form of Fourier dihedral were applied

to the related dihedral groups. Improper dihedrals defining two planes based on the groups of PSLG repeat unit were also included in the amino acid residue database.

The repeat units at both ends of the PSLG were not defined in the Gromacs termini database but in the amino acid residue database for simplicity. The N-termini has a total charge of 1; the middle repeat unit has a total charge of 0; the C-termini has a total charge of -1. Therefore, no matter how many repeat units a PSLG polymer has in the model, the total charge of the polymer is 0. The hydrogen database in Gromacs contains information for the pdb2gmx program on how to connect hydrogen atoms to existing atoms and needs to be updated. So each hydrogen in the repeat unit of PSLG was recorded following the format of the Gromacs hydrogen database.

Last, with the build-in command pdb2gmx, a topology file saving pairs, angles and dihedrals for PSLG could be generated if a coordinate file such as .pdb file was given in the right format.

3.2.2 Coarse-grained (Martini) parameterization of PSLG and THF

The first step of building a topology of THF for MD simulation is therefore to split the molecule in reasonable blocks. Like benzene, THF is a ring organic molecule, which is usually represented by three beads in the Martini FF approach. In order to preserve the geometry of small ring compounds, a four-to-one mapping procedure is inadequate. The strategy to model rings is therefore to include as many CG sites as necessary in order to keep the ring geometry, typically resulting in a 3 to 1 mapping of ring atoms onto CG beads [115]. With this more detailed mapping, enough geometrical detail is kept to mimic the geometry of small compounds such as THF or benzene.

During the THF modeling, three Martini beads, which belonged to two beadtypes, were represented three arbitrary groups a THF molecule. All three strategies had: 1. two apolar ring beads (Martini beadtype SC2) for carbon rings excluding oxygen; 2. one intermediate polar ring bead (Martini beadtype SN0) with oxygen. Three CG strategies had been investigated, as illustrated in Figure 35, but the first two were defective: the modeled THF beads packed too close, freezing the whole simulation box. Though the last did not cause the box freeze, the free energy results still had unacceptable large deviations from the experimental and AA simulation results.

Thus, a new THF model was built based on the third strategy, which is displayed in Figure 36. To improve the mobility of the system, the overall intramolecular "bond" interaction was reduced by increasing the "bond" length. The word of "bond" was in

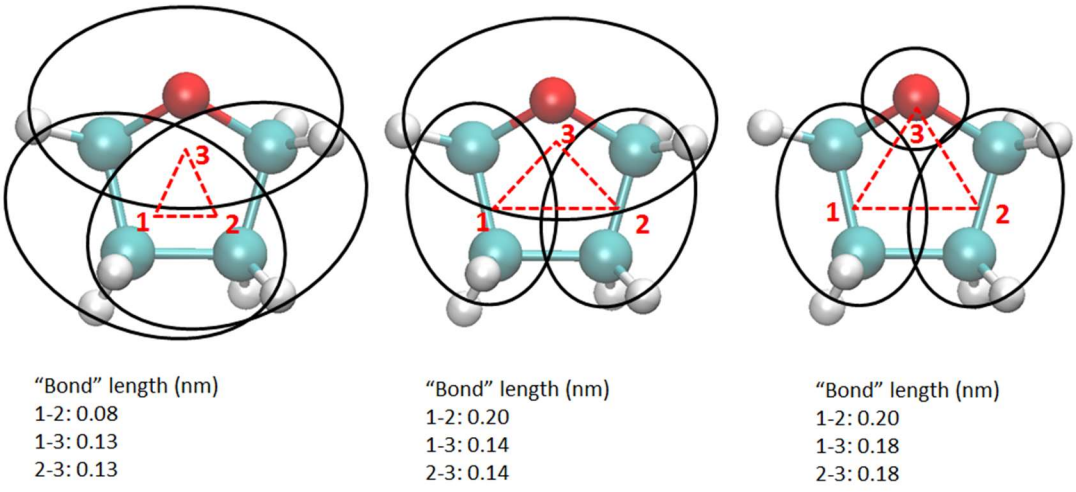


Figure 35. Three coarse-graining strategies of the THF AA trajectory. The vertex of the red-dash triangle is the center of mass of the arbitrary group (solid black eclipse). Red-dash lines are the arbitrary "bond" length in Martini, which equates the distances between centers of mass of each group. In each AA trajectory, cyan beads represent carbon; red bead represents oxygen; white beads represent hydrogen.

quotes because strictly speaking, no bond was defined for THF modeling with the Martini FF. Instead, the distance between the Martini beads were constrained. The "bond" length had been

increased 1.5 times longer than the distance of the center of mass of each AA group. LINCS constraint [102] was used to keep the bond length. Also, to further reduce the intramolecular bond interaction and improve the mobility of the system, the σ of the LJ potential was set to 0.36 rather than the 0.43 nm in the original Martini ring type, and the ε was scaled to 60% of the original value (75% in original Martini ring type).

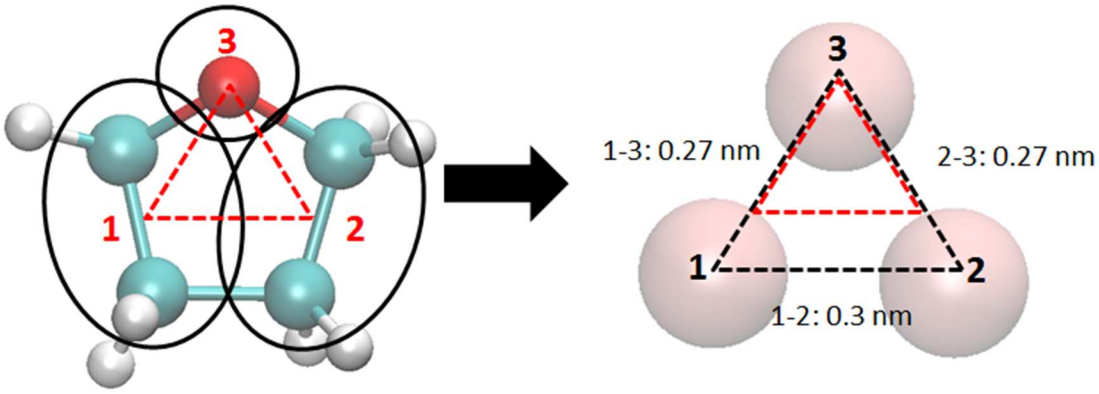


Figure 36. Schematic for the increasing of the "bond" length based on the center of mass distance of THF groups (red-dash lines). The new "bond" lengths (black-dash line) were 1.5 times longer than the distance of the center of mass of each AA group.

After the THF topology is available, the next step is to build a CG topology of PSLG. A PSLG repeat unit has several functional groups, so the way coarse-graining the groups is important. Therefore, similarly to the THF modeling, several strategies of group coarse-graining had been proposed and carefully examined (The selection details are not included in this paper for simplicity), and the chosen one is displayed in Figure 37. If the repeat unit is in the middle of the polymer, a nonpolar bead (Martini beadtype Nda) with a hydrogen-bond donor and an acceptor was assigned to represent the amide group, otherwise a strong polar bead (Martini beadtype P5) was assigned to represent the amide group. A nonpolar bead with an acceptor (Martini beadtype Na) was assigned to represent the ester group, while four apolar beads were assigned to represent the long alkene chain.

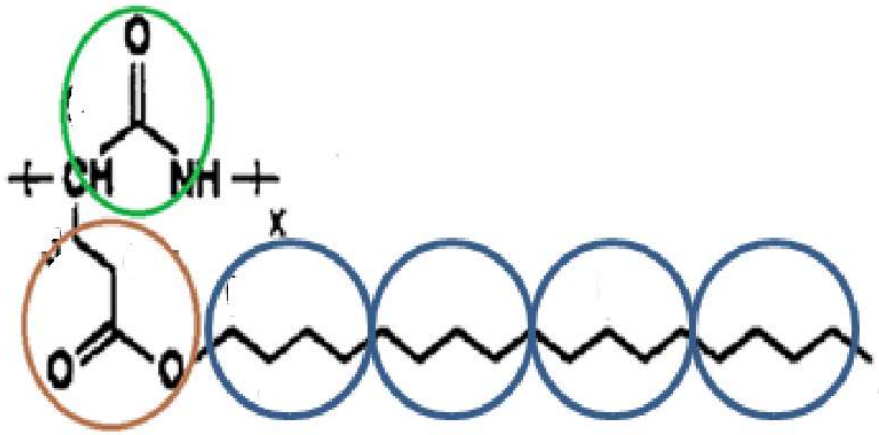


Figure 37. Coarse-graining strategy of the AA trajectory of a repeat unit of PSLG in the middle of polymer. Green circle: neutrally charged bead with a hydrogen-bond donor and an acceptor. Brown circle: neutrally charged bead with an acceptor. Blue beads: apolar. If the repeat unit is at the end of polymer, the head group in green would become strong polar.

Besides the bead assignment, the bonded interaction parameters was determined and tuned. The most appropriate choice of bonded parameters for the PSLG polymer was assessed from a comparison among CG, AA simulation results and experimental measurements of the RMSD, bond length and angel distributions. The final bonded parameters are listed in Table 8.

Table 8. Bonded parameters of PSLG with Martini Force Field

Type	Position	Length (nm)	Degree	Force constant ($kJ/nm^2 \cdot mol$)
bond	backbone-backbone	0.300	N/A	5000
bond	backbone-sidechain	0.378	N/A	7500
bond	sidechain-sidechain	0.495	N/A	2500
angle	backbone-backbone	N/A	77	1000
angle	backbone-sidechain	N/A	136	25
angle	sidechain-sidechain1	N/A	142	25
angle	sidechain-sidechain2	N/A	139	25

3.3 Method

3.3.1 Atomistic simulations

All molecular simulations were performed with the MD package GROMACS, version 4.6.7 [83]. Conventional MD simulation and thermodynamic integration were carried out separately. PSLG polymers with various repeat units ranging from 3 to 240 were created. As shown in Figure 38, a 15-repeat-unit PSLG has an alpha helix backbone and a long carbohydrate tail on each repeat unit. At first we planned to perform AA simulation for PSLG polymer having over 100 repeat units, but then found that AA simulations for larger PSLG polymer were going to be prohibitively expensive because the system for oil/water interface had a much larger number of molecules and thus the AA calculations of bonded and nonbonded molecular interactions require unaffordable simulation times. Calculations with CG models were performed in an attempt to overcome these limitations. Therefore, the goal of AA simulations was mainly to make comparisons with the experimental and CG results.

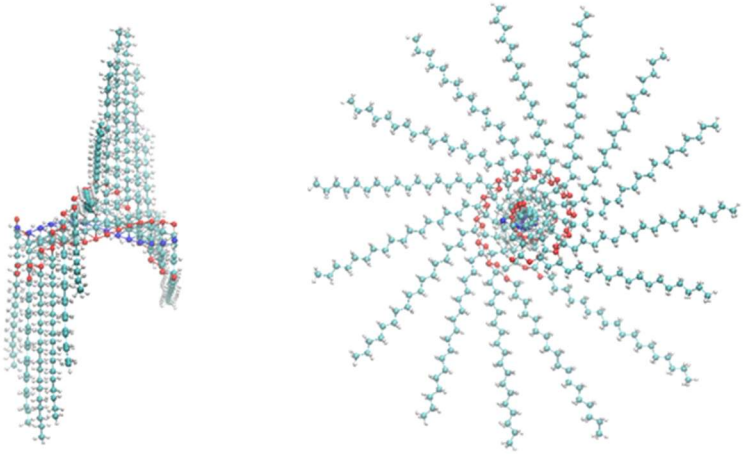


Figure 38. Atomistic molecular model of PSLG (repeat units = 15). Left, side view, amino terminus to the left; right, amino terminus end view. Cyan beads represent carbon; dark blue beads represent nitrogen; red beads represent oxygen; white beads represent hydrogen.

In the conventional MD simulation, a pure THF box with 1000 molecules was used for a 10ns simulation to test the mobility of the molecule, as well as the equilibrium density of the THF at 298K and 1bar. Then a PSLG polymer was initially in THF under periodic boundary conditions (PBC). The number of repeat unit varied from 3 to 15 with an increment of 3 in each box, and the backbone axis is parallel to the z-direction. Because of the length difference of the PSLG backbone, the size of each unit box varied slightly in z-direction (6 - 7 nm in x and y dimensions, while about 6 in z dimension if EAS was in single phase;), but all large enough to disallow direct interactions between periodic images. As shown in Table 9, System 1 to 5, the box was solvated with various number of THF molecules. To build the starting structure for the production run, firstly, a molecule of PSLG was centered in a rectangular box that had a minimum distance of 1.5 nm between any atoms of PSLG and the edge of the box.

Table 9. A summary of AA and CG conventional MD simulation of PSLG with small number of repeat units

System	Force field	PSLG length	THF molecule	Box size (nm)	Simulation Time (ns)
1	OPLS-AA	3	4536	8.8*8.8*8.8	20
2	OPLS-AA	6	6807	9.0*9.0*12.5	20
3	OPLS-AA	9	7078	9.0*9.0*13.0	20
4	OPLS-AA	12	7317	9.0*9.0*13.5	20
5	OPLS-AA	15	7572	9.0*9.0*14.0	20
6	Martini	3	4237	8.8*8.8*8.8	20
7	Martini	6	6193	9.0*9.0*12.0	20
8	Martini	9	6680	9.0*9.0*13.0	20
9	Martini	12	6919	9.0*9.0*13.5	20
10	Martini	15	7191	9.0*9.0*14.0	20

In the thermodynamic integration simulation, the free energy of vaporization and solvation change for the THF system was determined afterward by an appropriate numerical integration calculation [109, 110, 150-152]. The simulation gave the results of free energy of vaporization as well as free energy of hydration by changing a THF molecule's solvation state

from liquid to vacuum. The simulation strategy (i.e. simulation method, calculation) was the same as a study introduced by Liyana-Arachchi *et al*. [109, 110, 153]. The target THF molecule was first modeled under vacuum conditions by switching off all the interactions with the solvent molecules (THF or water). These interactions were then gradually turned on by coupling a parameter λ to the Hamiltonian H, where λ ranges from 0 (initial state, under vacuum) to 1 (final state, fully hydrated/solvated) with a step of $\Delta\lambda = 0.05$, which leads to 21 intermediate states with equally spaced values of λ. The free energies of vaporization and hydration were respectively determined by numerically integrating $\langle \partial H/\partial\lambda \rangle$ over the 21 simulations. In other words, to obtain the free energy of vaporization, a THF molecule was solvated in a pure THF box, and then the solvation state of that THF molecule was changed from pure THF liquid to vacuum by tuning the intermediate value of λ in the simulation environmental setting. Similarly, to obtain the free energy of vaporization, a THF molecule was solvated in a pure water box, and then the solvation state of the THF molecule was changed from water to vacuum by tuning the intermediate value of λ in the simulation environmental setting.

Before the production run, both conventional MD simulation and thermodynamic integration boxes were energy minimized with a steep descent method, and equilibrated in NVT ensemble at 298 K for 100 ps. Then molecules were further relaxed with 100 ps NPT MD simulation at 298 K and 1 bar. No barostat was applied if the system contained any gas phase. Position restraints of $1000 \cdot kJ \cdot mol^{-1}\ nm^{-2}$ were applied to every atoms of PSLG except hydrogen during the equilibrium step. Temperature was maintained at the target value using v-rescale coupling algorithm [99] with time constant $\tau_t = 0.1$, while pressure was maintained at the target value using Parrinello-Rahman coupling algorithm [100] with time constant $\tau_p = 3.0$. Bonded length were constrained using the LINCS algorithm [102] for other components. The Lennard-

Jones interaction was truncated at the distance of 1.1 nm. The particle-mesh Ewald (PME) method [103] with a cutoff of 1.1 nm and a grid spacing of 0.16 nm was used for taking account of all columbic interaction, where the reaction-field method was used [104, 105]. The long range dispersion corrections were applied for energy and pressure. Finally, a 20 ns production run with time step of 1 fs was applied for each system in the conventional MD simulations. For the simulation of thermodynamic integration, the simulation time was 2.5 ns for each transition state, of which at least the last 2 ns were used to calculate the averages.

3.3.2 Coarse-grained simulations

We considered simulations with CG models to investigate systems of interest with large dimensions for longer time scales. Conventional MD simulation and thermodynamic integration were carried out separately. The first goal of CG simulations was to validate the CG model of by comparing the CG results with the experimental and AA results. Then the CG model of PSLG was validated by assessing CG simulation results, AA simulation results and experimental measurements of the RMSD, bond length and angel distributions. Last, hydrophobins were coated on an n-decane/water interface, while several PSLG polymers were in the n-decane. The orientations of the PSLG polymers were investigated.

We used the Martini force field [114-117] for PSLG, THF, hydrophobin, n-decane, water and NaCl. In the conventional MD simulation, to verify and optimize the nonbonded parameters of THF, a pure THF box with 1000 molecules (3000 beads, as it was 3 to 1 mapping) was run for 10 ns to test the mobility of the molecule, as well as the equilibrium density of the THF at 298K and 1bar. Same to the AA simulation of one THF molecule in the THF or water box, in the thermodynamic integration simulation, the total free energy change for the THF system was determined afterward by an appropriate numerical integration procedure. The simulation gave

the results of free energy of vaporization and free energy of hydration by changing a THF molecule's solvation state from liquid to vacuum. To obtain the free energy of vaporization, a THF molecule was solvated in a pure THF box, and then the solvation state of that THF molecule was changed from pure THF liquid to vacuum by tuning the intermediate value of λ in the simulation environmental setting. Similarly, to obtain the free energy of vaporization, a THF molecule was solvated in a pure water box, and then the solvation state of the THF molecule was changed from water to vacuum by tuning the intermediate value of λ in the simulation environmental setting. The mobility and density of a THF box was also tested through the convention simulation in order to evaluate nonbonded parameters.

In the PSLG model testing, as the topology of a single PSLG repeat unit was made available, PSLG polymers with various repeat unit number ranging from 3 to 240 were created respectively. Figure 39 illustrates a schematic for the mapping of an AA structure of PSLG

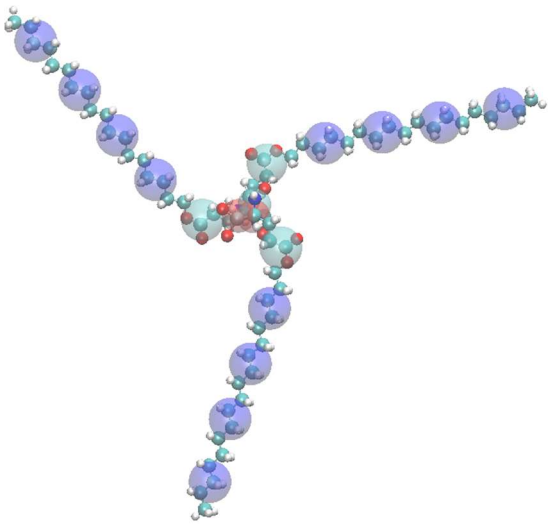


Figure 39. Schematic for the mapping of an AA structure of PSLG (repeat unit = 3) to a CG structure, conducted by the Martini force field. In the AA system (small beads), cyan beads represent carbon; dark blue beads represent nitrogen; red beads represent oxygen; white beads represent hydrogen. In the CG system (large semitransparent beads), the amides forming polymer backbone are represented as red beads; the ester groups are represented as cyan beads; the long alkene chains are represented as dark blue beads.

molecule to a CG structure. Though the long carbon tail in each repeat unit looked very straight, the polymer was actually semi-flexible. Therefore, the parameters were made to bend the long carbon chain to some extent right after the beginning of the simulation. After the coarse-graining process, there were fewer sites in the system. Thus, both inter- and intra-molecular interactions among the sites were greatly reduced. A modified Martini Version 2.2 P is used. The parameter modification details could be found in section 4.2.2 when modeling the CG topology of THF.

Table 9 shows systems (6 through 10) for the CG conventional MD simulations, surrounding environment of which were comparable to the AA conventional MD simulations. The scenarios for preparing each box of System 6 to System 9 were the same as those used in AA simulations.

By conducing both AA and CG simulations, we are able to compare many physical properties obtained through AA and CG simulations, then we can use this "partial verified" CG method to probe more complicated systems that are too large to be modeled by AA force fields. As shown in Figure 40, a PSLG polymer having 240 repeat units was modeled, size of which is

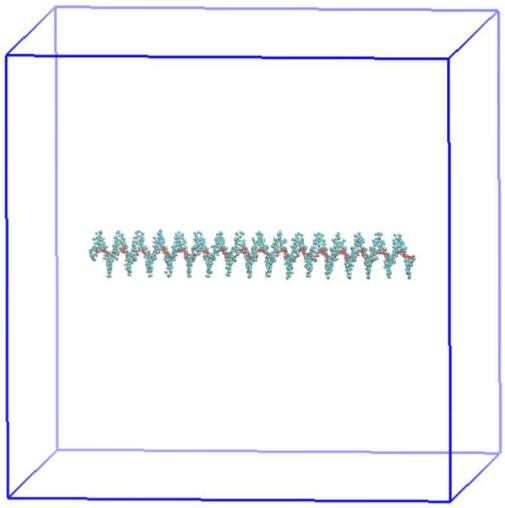


Figure 40. The simulation box of PSLG with 240 repeat units in THF. The simulation box has a dimension of 50 nm * 50 nm * 50 nm. The molecule of PSLG is represented by red beads (for backbone) and cyan beads (for sidechains). For clarity, THF is not shown.

comparable to the available experimental results. The simulation box had a dimension of 50 nm * 50 nm * 50 nm, which was large enough to prevent direct interactions between periodic images. Once the box had been created, 878477 THF molecules were added to solvate the PSLG molecule. We measured PSLG's size and shape, and compared to the experimental results when available.

Finally, with this "total verified" CG model of PSLG more complicated system including the interface between multiple PSLG molecules and the hydrophobins encapsulating them was created to test alignment and packing of semiflexible polymers by hydrophobins and to suggest ways to optimize the capture and alignment of the polymers in these systems. Figure 41 is the schematic for the illustration of the simulation boxes. To build the starting structure for the production run, firstly, either 6 ("low" PLSG concentration) or 9 ("high" PLSG concentration) molecules of PSLG were centered in a rectangular box that had a minimum distance of 5 nm between any beads of PSLG and the edge of the box. Each molecule of PSLG was parallel to each other. By rotating the PSLG molecules, two PSLG alignments of interest were prepared: (1) backbones of all PSLG were perpendicular to the hydrophobin coated oil/water interface; (2) backbones of all PSLG were parallel to the oil/water interface. The PSLG molecules were then solvated in n-decane or THF in a periodic simulation box of dimensions of x, y, and z, where dimension z corresponds to the backbone chain orientation if all PSLG were perpendicular to the hydrophobin, dimension y corresponds to the backbone chain orientation if all PSLG were parallel to the hydrophobin. The hydrophobin monolayer was at the plane of x and y.

To build the simulation box displayed in Figure 41, an extra pure water box as well as an extra hydrophobin monolayer box were created and equilibrated at 298K and 1bar, both of which having the same x and y dimensions (26 nm * 26 nm) as the n-decane (THF) box. Then these

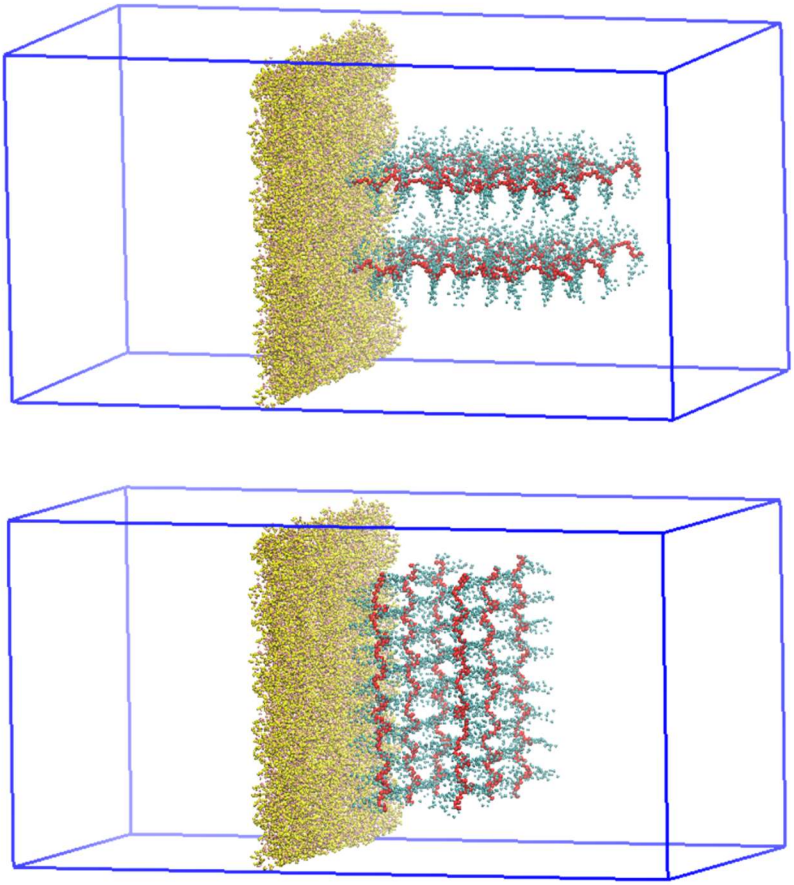


Figure 41. Schematic for the illustration of different simulation boxes. Upper: the backbones PSLG molecules are perpendicular to the EAS hydrophobin monolayer; bottom: the backbones of PSLG molecules are parallel to the EAS hydrophobin monolayer. The molecule of PSLG is represented by red beads (for backbone) and cyan beads (for sidechains). The molecule of EAS hydrophobin is represented by pink beads (for backbone) and yellow beads (for sidechains). PSLG was solvated in decane or THF on the right side. Water box was on the left side. For clarity, decane/THF and water molecules were omitted.

three boxes were integrated into one simulation box. The final integrated simulation box has a dimension of 26 nm *26 nm *52 nm, consisting 6 (“low” PLSG concentration) or 9 (“high” PSLG concentration) PSLG molecules, ~55000 decane or THF molecules, 120 EAS hydrophobins and 107632 water molecules. 240 Na+ ions was added to neutralize the system. Every PSLG molecules in the simulation box are identical. After careful consideration, the number of repeat unit in each PSLG molecule was set to 120 to make sure while computational

resource became affordable (even the simulation is carried on with CG model, a couple months is required for the desired simulation length), the size of the PSLG was still comparable to the minimal experimental detectable size.

The simulation parameters for computing Lennard-Jones interaction and Coulombic interaction were the same as those published in the original Martini papers [115, 116]. Before the production run, both conventional MD simulation and thermodynamic integration systems were energy minimized with a steep descent method, and equilibrated in NVT ensemble at 298 K for 500 ps. Then molecules were further relaxed with 5000 ps NPT MD simulation at 298 K and 1 bar. No barostat was applied if the system contained any gas phase. Position restraints of $1000 \cdot kJ \cdot mol^{-1}\ nm^{-2}$ were applied to every atoms of PSLG except hydrogen during the equilibrium step. Temperature was maintained at the target value using v-rescale coupling algorithm [99] with time constant $\tau_t = 1$, while pressure was maintained at the target value using Parrinello-Rahman coupling algorithm [100] with time constant $\tau_p = 8.0$. Bonded length were constrained using the LINCS algorithm [102] for other components. Once the system was energy-minimized and well equilibrated, production runs in the THF and PSLG force field validation were carried out with a time step of 10 fs and a total simulation time of 20 ns. The large box for the PSLG with 240 repeat units was run for 40ns. The large model of the interface between the semiflexible PSLG and the hydrophobins was run for 500-1000ns.

3.4 Results and discussion

3.4.1 Validation of CG force field for THF model

To verify the CG force field for the THF model, both conventional and thermodynamic integration simulations were carried out. The former was to examine if the CG model of THF can reproduce some physical properties such as bond length and angle distributions of THF,

while the later was to measure the free energy properties and make a comparison with the results from AA simulation and experiment. After tuning the parameters of THF multiple times, an optimized CG THF model was finally built.

Table 10 compares some physical properties obtained from both AA and CG simulations of a pure THF box, as well as the experimental results. Both AA and CG simulations gave similar results for densities that are very close to the reported experimental density value of 883 kg/m3 at 298K and 1 bar [154]. Also, we observed the mobility of THF beads in the simulation box through VMD [33] to make sure the system was not freezing. We found both THF molecules in AA or CG boxes had certain mobility, which agrees with THF's liquid state at 298K and 1 bar. Nonbonded parameters were optimized on the basis of comparison of densities of THF to experimental values as well as comparison of mobility of single THF molecule and whole THF box to mapped atomistic simulation results.

We calculated the experimental value of the free energy of vaporization according to the Gibbs free energy equation at constant temperature:

$$\Delta G = \Delta H - T\Delta S \tag{7}$$

From the reported experimental values of the heat [155] and entropy of vaporization [156, 157] for THF, its free energy of vaporization is about 2.46 kJ/mol. The simulated free energy of evaporation was found to be 16.34 for the AA model, while the simulated free energy of vaporization was found to be 18.35 kJ/mol. Though close to each other, both results from AA and CG models have significant deviations from the calculated experimental value. Nevertheless, the measurement of free energy of evaporation was less important because it was only a way to compare if the AA and CG results were close enough or not. The following simulations did not require an accurate result for the free energy of vaporization.

Likewise, the AA simulated free energy of hydration for THF in water was calculated using thermodynamic integration, which was equal to 14.22 kJ/mol, while the CG simulated free energy of hydration for THF in water was equal to 13.33 kJ/mol. These values are reasonably close to the reported experimental free energy of hydration for THF (14.51 kJ/mol) [158, 159]. This favorable comparison between simulation and experimental indicate that our choice of molecular models can reproduce physical properties of these systems.

Table 10. A physical and thermodynamic property summary of conventional MD and thermodynamic integration simulations for the THF force field validation.

	Density (kg/m3)	Freezing	Gvap (kJ/mol)	Ghydra (kJ/mol)
OPLS-AA	871	No	16.34	14.22
Martini-CG	860	No	18.35	13.33
Experiment	883[a]	No	2.80[b]	14.51[c]

Gvap is the free energy of vaporization; Ghydra is the free energy of hydration.
a Experimental value of density is from [154].
b Experimental value of Gvap is calculated from the heat of vaporization and entropy of vaporization [155-157].
c Experimental value of Ghydra is from [158, 159].

Coarse-grained parameters for the bond and angle potentials were optimized by comparing distributions with AA simulations of THF mapping them to the CG representation. Each THF molecule has 3 groups in AA trajectory or 3 beads in the CG trajectory, and then the bond length is defined as the distance between the centers of the group (bead), while the angle is the space between the two bonds. Shown in Figure 42, AA and CG bonded distributions were plotted as a function of bond length or angle. Gromacs 4.6.7 [83] offers useful tools (g_bond and g_angle) to extract the relevant bonded distributions of the groups in AA simulations beads in CG simulations. The green lines in Figure 42, obtained from CG simulation agree well with the all atom distributions, demonstrating the success of the parametrization of the bond terms. The polar-apolar bond length of AA and CG both had max probabilities around 0.18 nm, while the

apolar-apolar bond length of AA and CG both had max probabilities around 0.2nm, indicating an acceptable agreement with each other. Though the peak of CG in the apolar-apolar bond length distribution was lower than the AA, the overall distribution widths of AA and CG were still approximately the same.

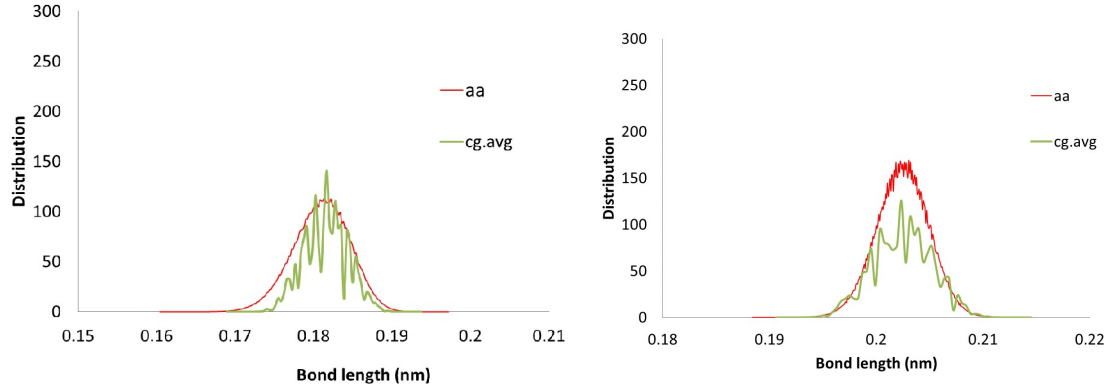


Figure 42. Comparison of AA and CG bond length distributions for THF. Left: distribution of the polar-apolar groups (beads). Right: distribution of the apolar-apolar groups (beads). Red line is obtained from AA simulation; green line is obtained from CG simulation averaging every 5 bins.

Also, the blue line in Figure 43, obtained from CG simulation agrees well with the all atom distributions, demonstrating the success of the parametrization of the angle terms. The polar-apolar-apolar angles of AA and CG both had max probabilities around 56 degree, while the apolar-polar-apolar angles of AA and CG both had max probabilities around 68 degree, implying a good agreement with each other. Though the peaks of CG were 20% higher than the AA and the overall distributions of CG were narrower than the AA, indicating lower force constants should be applied to make the CG angle distribution more consistent with AA, the deviation is acceptable and the applied force constants in the CG model are considered optimized. This is because changing force constants in the CG model would significantly change other properties such as bond length distribution. The eventual force constants were a compromise and therefore considered as overall optimal.

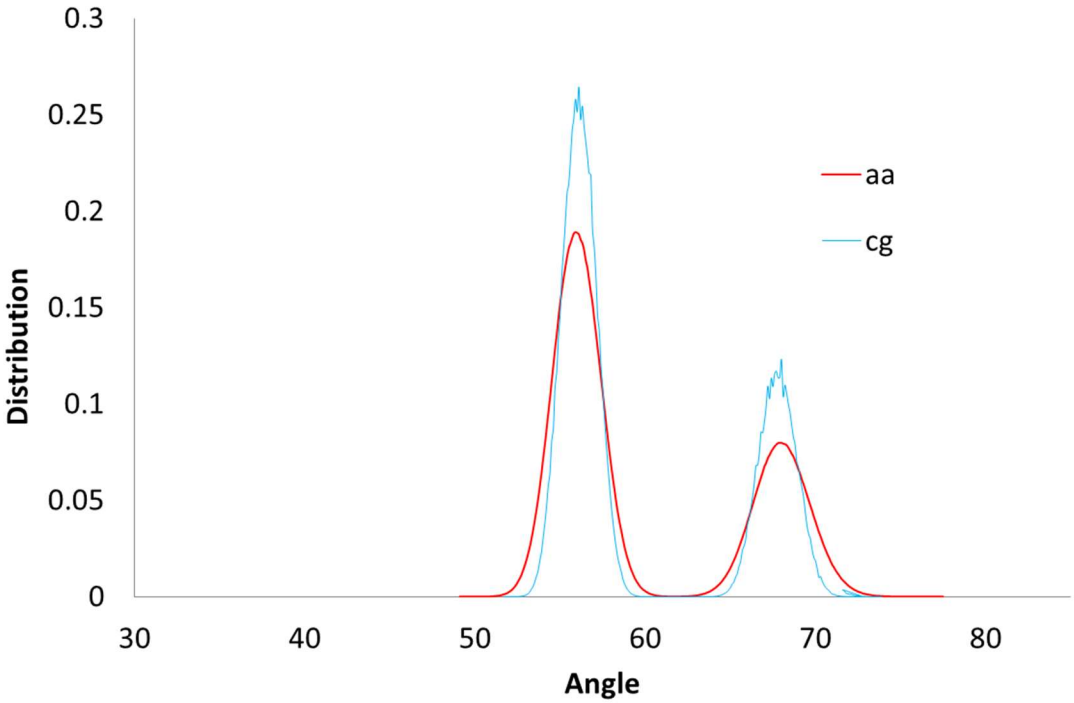


Figure 43. Comparison of AA and CG angle distributions for THF. The first distribution peak appearing at 56 degree is the angle of polar-apolar-apolar group (beads), the second distribution peak appearing at 68 degree is the angle of apolar-polar-apolar group (beads). Red line is the obtained from AA simulation; blue line is obtained from CG simulation.

3.4.2 Validation of AA and CG force fields for PSLG model

The most appropriate choice of bonded and nonbonded parameters for the PSLG polymer was assessed from a comparison among CG and AA simulation results, as well as with several experimental measurements (RMSD, bond length and angle distributions). Therefore, similar to the THF validation tests, the PSLG model validation process also included bonded distribution measurements of several relevant groups, and comparison between AA and CG results. The validation process of PSLG models, however, was much more difficult than the THF. Unlike small organic solvent such as THF, each repeat unit had 72 atoms (compared to 13 atoms in a THF AA molecule) or was represented by 6 beads (compared to 3 beads in a THF CG molecule). Hence, each polymer molecule has various repeat units (up to 240 in this work), and the

correlations among the repeat united must be considered. Both of the AA and CG models have to reproduce properties of polymers with different repeat units as much as possible.

We first built a PSLG polymer having 3 repeat units in AA trajectory with the parameters introduced in Section 4.2.1, and then mapped them to the CG representation. The average bonded distributions between the backbone groups in AA trajectory or beads in CG trajectory were examined at first, as we thought the overall structure of a polymer was greatly influenced by the backbone connections. It is worth to mention that the interaction between the two backbone beads of terminus (both assigned a P5 beadtype) was strong and had a great impact on the bonded distribution, so during the CG parameter optimization process, the 12−6 LJ parameter σ was changed to 0.36 to make the CG results more close to those from AA simulations. This modification was unique and only for the PSLG with 3 repeat units, because when longer chain of PSLG was modeled, the end to end distance increased so interactions between two terminuses became insignificant. As illustrated in Figure 44, each repeat unit has a backbone group (bead), and each PSLG polymer with 3 repeat units has 3 backbone groups (beads). As a consequence, the bond length as showing in black arrows is defined as the distance between the centers of the group (bead) connected to one another. Thus, since a PSLG polymer with 3 repeat units had two bonds, the averaged backbone-backbone bond length distribution had been checked and plotted in Figure 44. The green lines in Figure 44, obtained from CG simulation agree well with the AA results. The backbone-backbone bond lengths from AA and CG simulations both had max probabilities around 0.32 nm. Even though the peak of CG in the backbone-backbone bond length distribution was slightly lower than that for AA results, and the CG distribution at the larger bond length zone was slightly wider than AA, the overall width of the AA and CG simulations were still approximately the same.

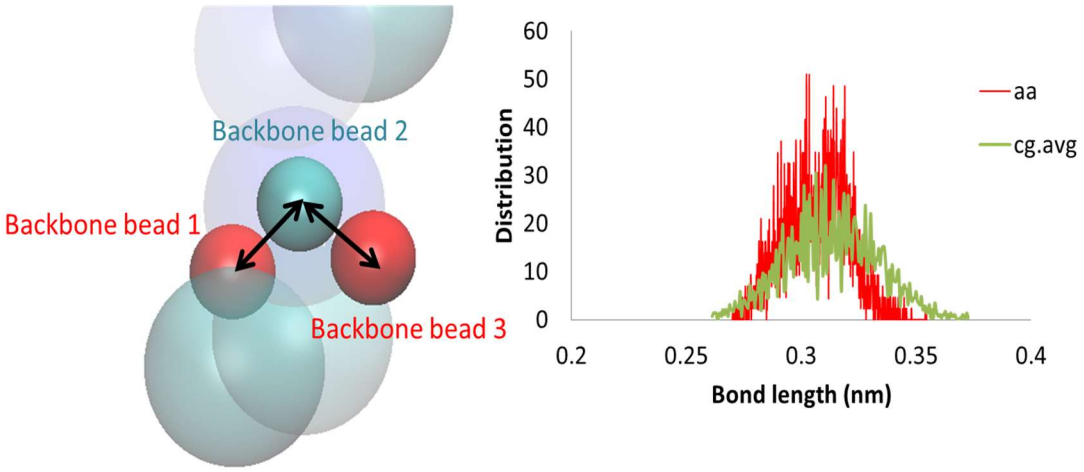


Figure 44. Comparison of AA and CG backbone-backbone bond length distributions for PSLG (repeat unit = 3). The left subfigure elucidates that each repeat unit has a backbone group (bead), and the bond length as showing in black arrows is defined as the distance between the centers of the group (bead) connecting to one another. In the right sub figure, red line was obtained from AA simulation; green line was obtained from CG simulation, which was smoothened by running averages from bond length of 0.5 nm width.

Next, we measured distributions of the angle between two bonds of the centers of the backbone group (bead) connected to one another. The obtained angle distribution from CG simulation was consistent with the distributions from AA simulations; as shown in Figure 45, the

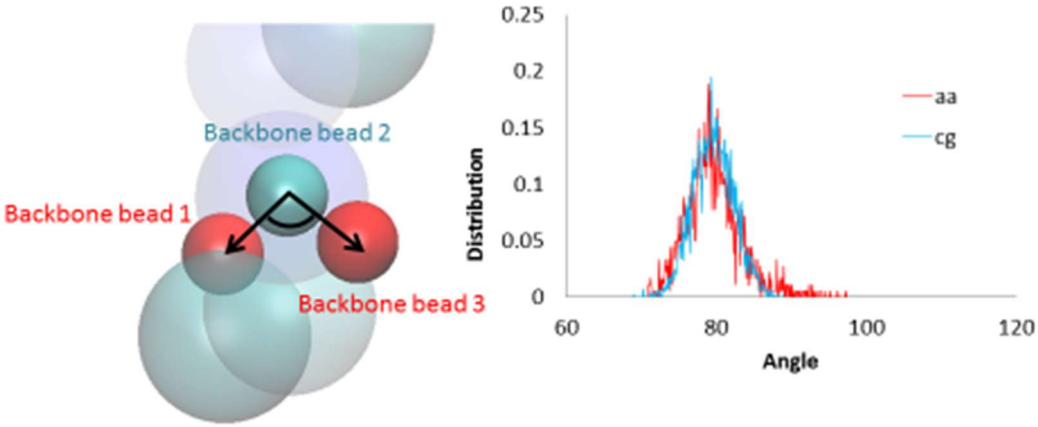


Figure 45. Comparison of AA and CG backbone angle distributions for PSLG (repeat unit = 3). The left subfigure elucidates that the PSLG polymer (repeat unit = 3) has a backbone angle, which is the space between two black arrows as the bonds of the centers of the backbone group (bead) connected to one another. In the right subfigure, the distribution peak appearing at 79 degree is the angle of two backbone groups (beads). Red line is the obtained from AA simulation; blue line is obtained from CG simulation.

backbone angles of AA and CG both had max probabilities around 79 degree, indicating a very good agreement with each other. The peak of CG was in good match of the AA and the overall distributions of CG were remarkably similar to AA. Besides the bonded distributions of the backbone groups (beads), we also measured bonded distributions of other groups (beads) in the PSLG polymer with 3 repeat units, and would introduce the details of these measurements along with the tests for PSLG polymer with more than 3 repeat units.

Next, PSLG polymers with repeat units = 6, 9, 12 and 15 were created respectively in both AA and CG trajectories. Figure 46 displays a schematic for the mapping of an AA structure of PSLG with 15 repeat units to a CG structure. The backbone had an α-helix structure which was expected to be preserved during the simulation. Though the long carbon tail in each repeat unit was totally straight in the initial structure, the polymer was actually semi-flexible so the parameters were tweaked to bend the long carbon tail to some extent soon after the simulation

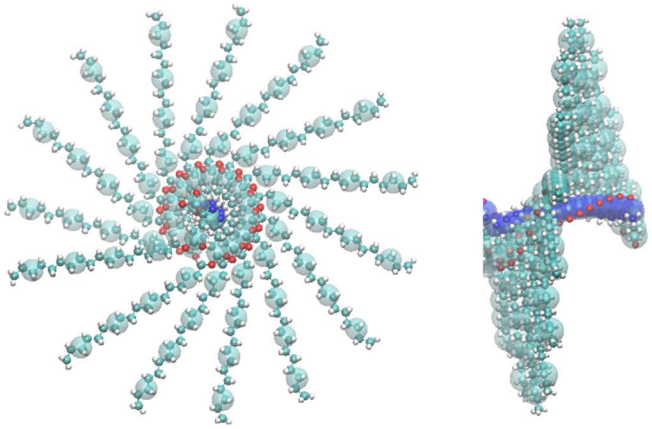


Figure 46. Schematic for the mapping of an AA structure of PSLG (repeat unit = 15) to a CG structure. Left: amino terminus end view; right: side view, amino terminus to the left. In the AA system (small beads), cyan beads represent carbon; dark blue beads represent nitrogen; red beads represent oxygen; white beads represent hydrogen. In the CG system (large semitransparent beads), the amides forming polymer backbone are represented as large blue beads; the ester groups and the long alkene chains are represented as large cyan beads.

started. Then the created PSLG polymers were solvated with THF first, followed by conventional MD simulations, and lastly the bonded distributions for each PSLG were measured.

At first, distributions of three kinds of bond length, i.e. backbone-backbone, backbone-sidechain and sidechain-sidechain, for each PSLG polymers having repeat units from 3 to 15 were probed. All of the three bond length distributions are recorded from Figure 48 to Figure 50.

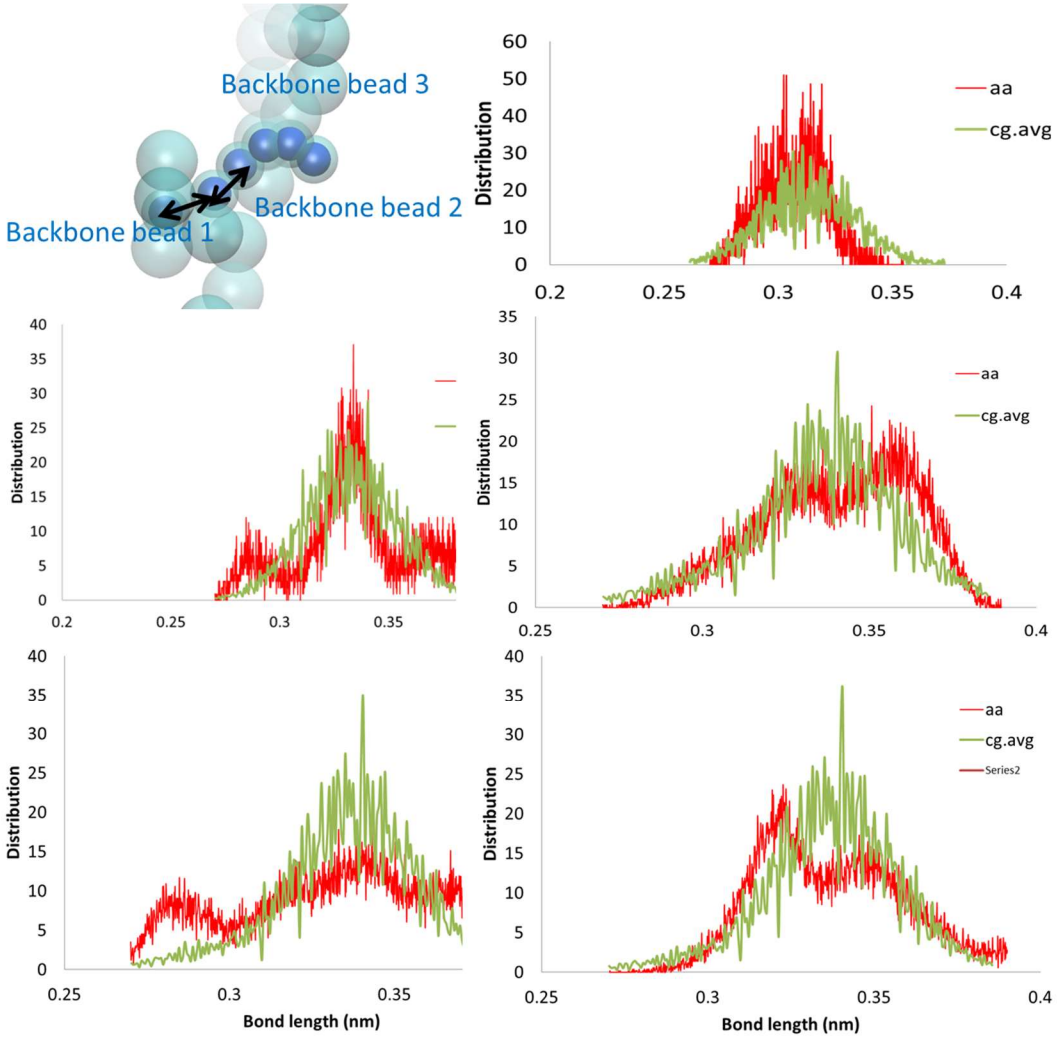


Figure 47. Comparison of AA and CG backbone-backbone bond length distributions for PSLG (repeat unit = 3, 6, 9, 12, 15 displayed in the order of upper right, middle left, middle right, bottom left and bottom right). The upper left subfigure elucidates that each repeat unit has a backbone group (bead), and the bond length as showing in black arrows is defined as the distance between the centers of the group (bead) connecting to one another. In the rest subfigures, red line was obtained from AA simulation; green line was obtained from CG simulation, which was smoothened by running averages from bond length of 0.5 nm width.

The backbone-backbone bond length is defined as the distance between the centers of the groups (bead) connecting to one another (only the repeat units next to each other has backbone-

backbone bond); the backbone-sidechain bond length is defined as the distance between the centers of the backbone group (bead) and the first sidechain group (bead) in each repeat unit. The distance between the centers of the sidechain groups (beads) is referred as the sidechain-sidechain bond length. In Figure 48, we found both backbone-backbone bond lengths of AA and CG in all 5 PSLG polymers except repeat unit =3 had max probabilities around 0.34 nm. In the

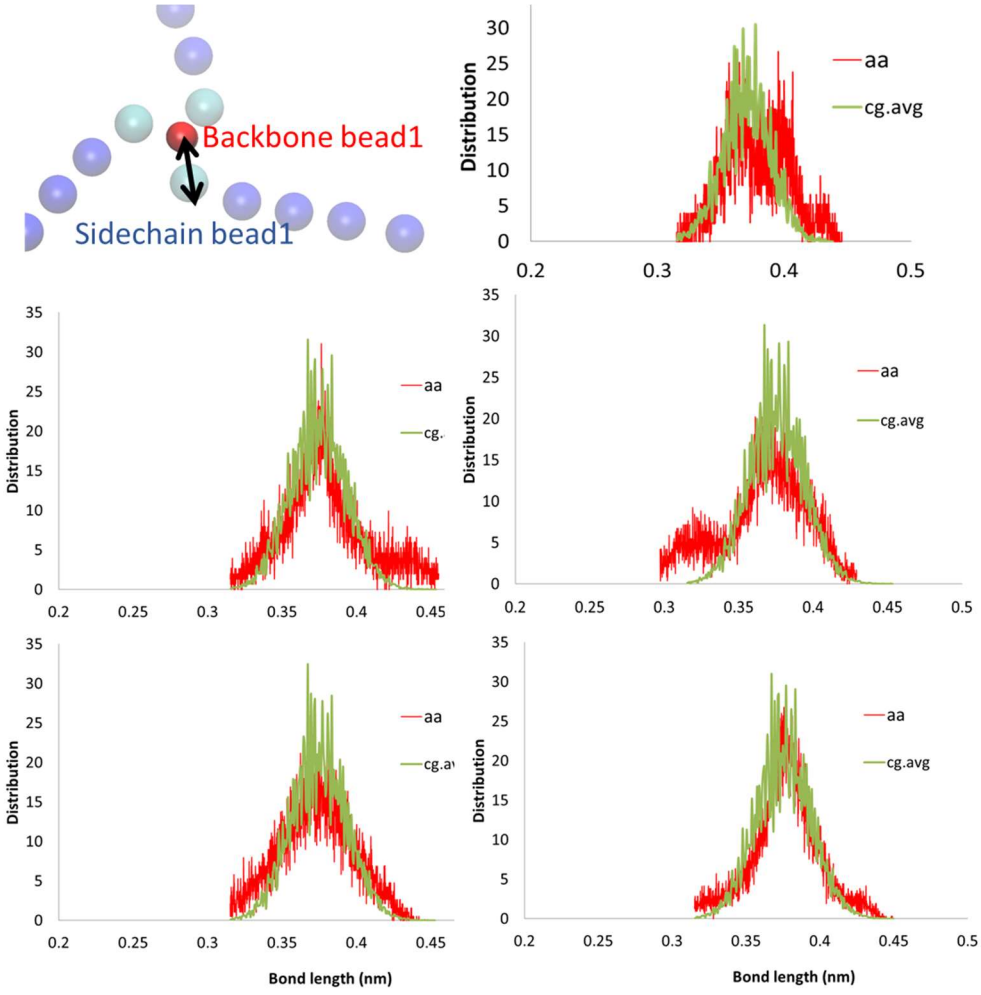


Figure 48. Comparison of AA and CG backbone-sidechain bond length distributions for PSLG (repeat unit = 3, 6, 9, 12, 15 displayed in the order of upper right, middle left, middle right, bottom left and bottom right). The upper left subfigure elucidates that in each repeat unit the distance between the centers of the backbone group (bead) in red and the first sidechain group (bead) in cyan is the backbone-sidechain bond length showing in black arrow. In the rest subfigures, elucidation is the same as Figure 47.

repeat unit = 3 case, while its AA and CG distributions were close to each other, the backbone-backbone bond length was shorter possibly because of stronger interaction between the end terminuses. The strong interaction between the end terminuses in PSLG with 3 repeat units didn't affect the backbone-sidechain bond length. As a result, we found both backbone-sidechain bond lengths of AA and CG in all 5 PSLG polymers had max probabilities around 0.36 nm,

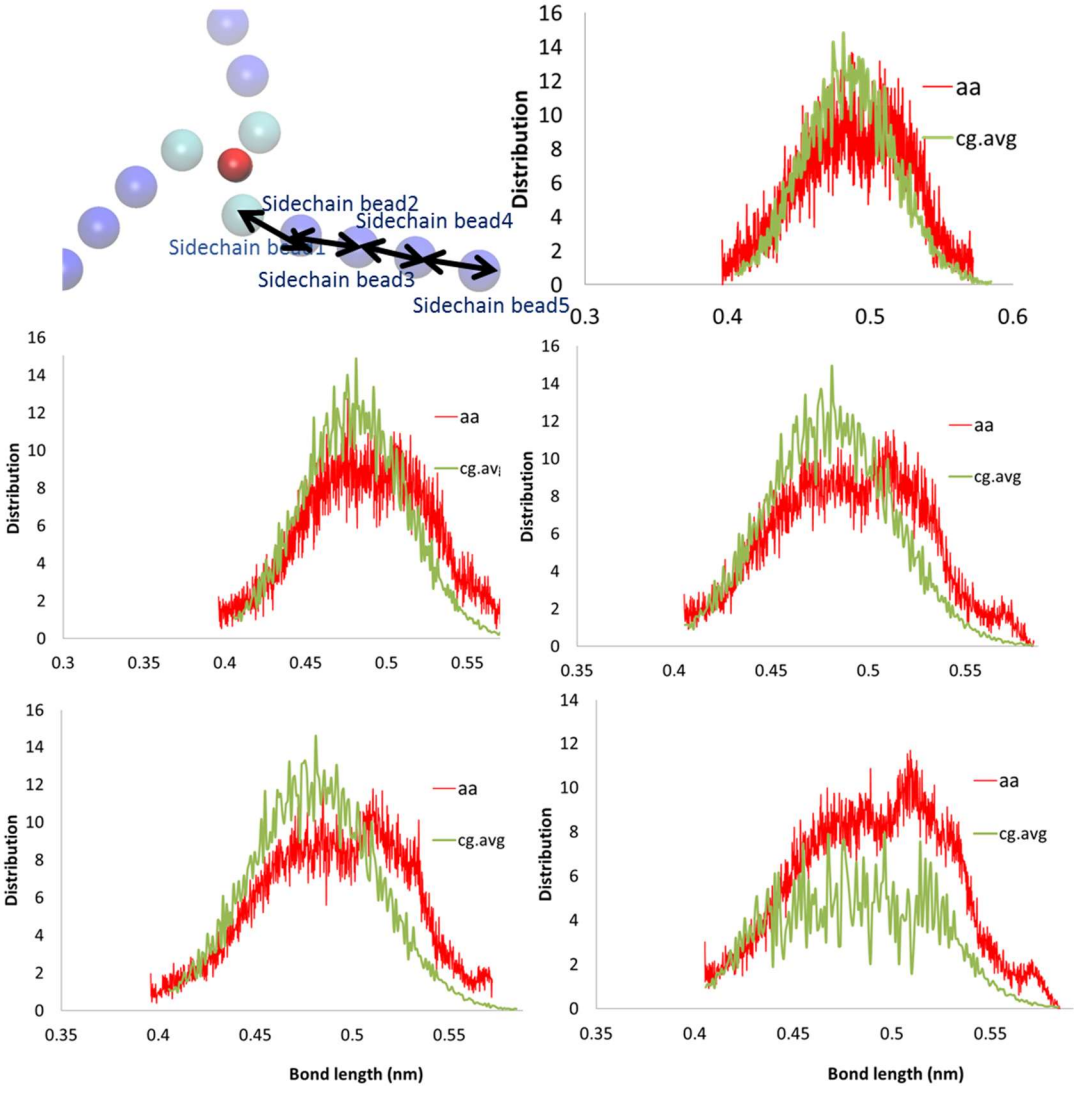


Figure 49. Comparison of AA and CG sidechain-sidechain bond length distributions for PSLG (repeat unit = 3, 6, 9, 12, 15 displayed in the order of upper right, middle left, middle right, bottom left and bottom right). The upper left subfigure elucidates that in each repeat unit the distance between the centers of the sidechain groups (beads) in either cyan and blue or blue and blue is the sidechain-sidechain bond length showing in black arrows. In the rest subfigures, elucidation is the same as Figure 48.

while sidechain-sidechain bond lengths of AA and CG in all 5 PSLG polymers had max probabilities around 0.47 nm.

Next, four kinds of angle distributions for each PSLG polymers having repeat units from 3 to 15 were probed. We categorized angles in PSLG polymer to backbone-backbone angle, backbone-sidechain angle and backbone-sidechain angle; the last had two different types because of the angle position difference. The backbone-backbone angle is the space between the bonds of the centers of the backbone group (bead) connected to one another; the backbone-sidechain angle is the space between the backbone-sidechain bond and sidechain-sidechain bond; the sidechain-sidechain angle is the space between the sidechain-sidechain bonds. All of the four angle distributions are recorded from Figure 50 to Figure 53. In Figure 50, we found both backbone-backbone angles of AA and CG in all 5 PSLG polymers except repeat unit =3 had max probabilities around 110 degree. Again in the repeat unit = 3 case, while its AA and CG distributions were close to each other, the backbone-backbone angle was shorter possibly because of stronger interactions between the end terminuses. The strong interaction between the end terminuses in PSLG with 3 repeat units didn't affect the backbone-sidechain angle. As a result, in Figure 51, we found both backbone-sidechain angles of AA and CG in all 5 PSLG polymers had max probabilities around 145 degree. The angle is not 180 degree (straight chain), this is because the PSLG molecules are semiflexible in many organic solvent such as THF. Therefore, the long carbon chains were expected to show a deviation from the starting structure as shown in Figure 46, and should not be the rigid straight (sidechain-sidechain angle = 180 degree). On the other hand, the carbon chains of PSLG should not be completely random as coils because of its semiflexibility observed via experiment. The angles were well constrained at a certain degree greater than 100 (140 degree for the sidechain-sidechain angle) in both of our AA

and CG models, and the distribution was neither too narrow (become rigid at that degree) nor too wide (constraint was too weak). Consequently, the distribution situation was in accordance with the experimental observation.

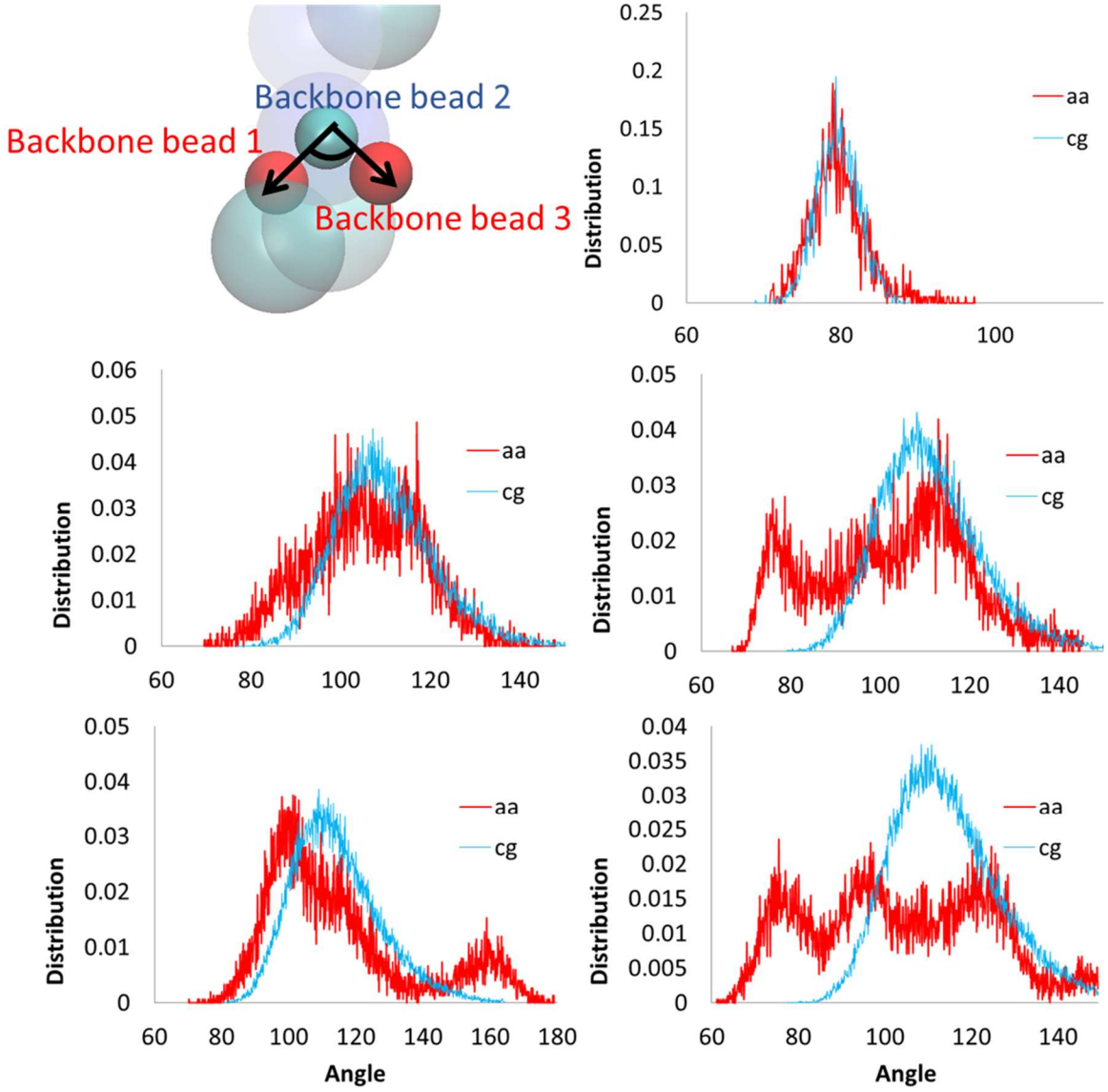


Figure 50. Comparison of AA and CG backbone-backbone angle distributions for PSLG (repeat unit = 3, 6, 9, 12, 15 displayed in the order of upper right, middle left, middle right, bottom left and bottom right). The upper left subfigure elucidates that the backbone-backbone angle is the space between the bonds (showing in black arrows) of the centers of the backbone group (bead) connected to one another. In the rest subfigures, red line is the obtained from AA simulation; blue line is obtained from CG simulation.

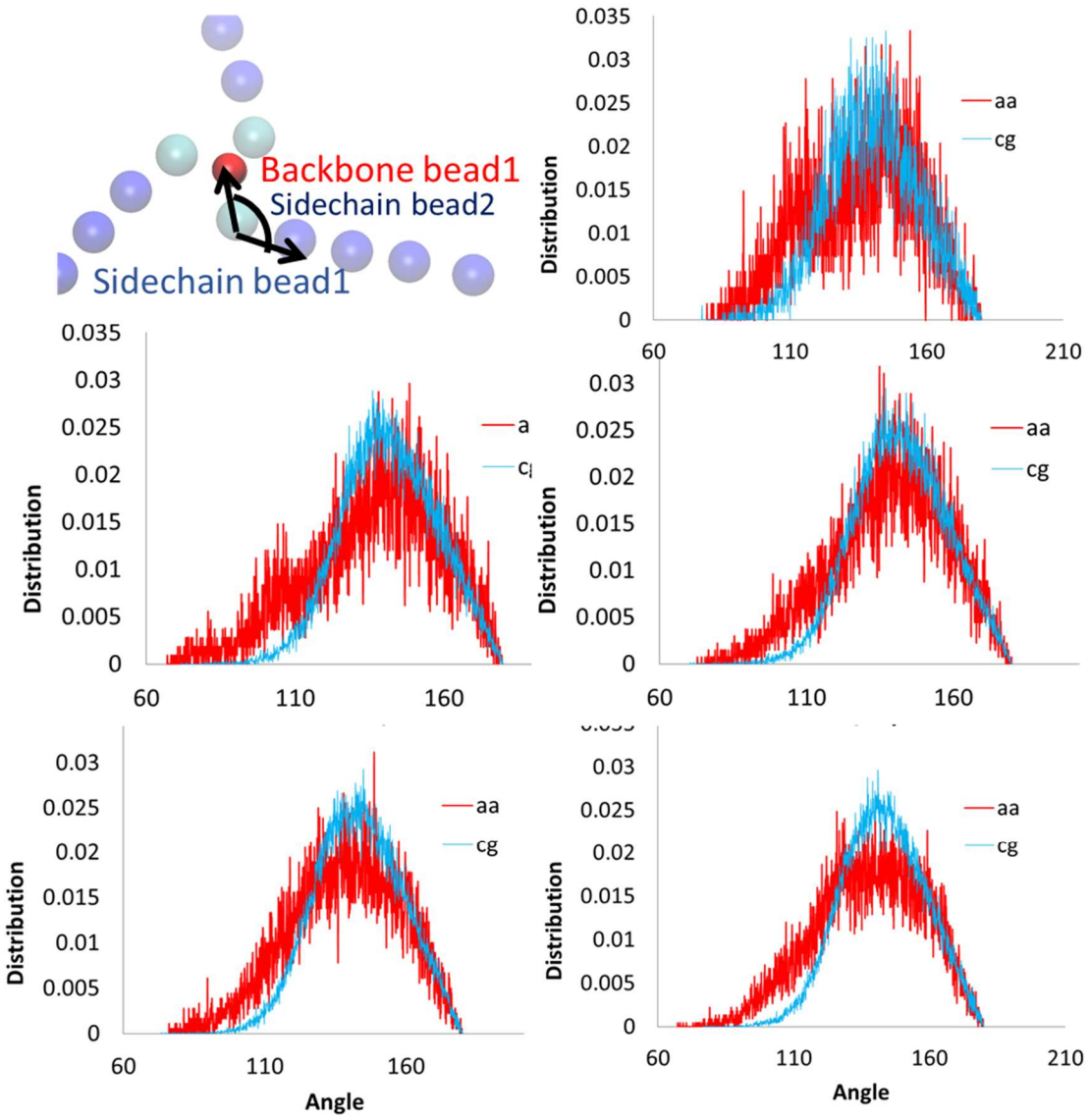


Figure 51. Comparison of AA and CG backbone-sidechain angle distributions for PSLG (repeat unit = 3, 6, 9, 12, 15 displayed in the order of upper right, middle left, middle right, bottom left and bottom right). The upper left subfigure elucidates that the backbone-sidechain angle is the space between the backbone-sidechain bond and sidechain-sidechain bond (showing in black arrows). In the rest subfigures, red line is the obtained from AA simulation; blue line is obtained from CG simulation.

In Figure 52 and Figure 53 sidechain-sidechain angles of AA and CG in all 5 PSLG polymers had max probabilities around 140 degree. Again, this is because the PSLG molecules are semiflexible in THF. Therefore, the long carbon chains were expected to show a deviation from the starting structure in model.

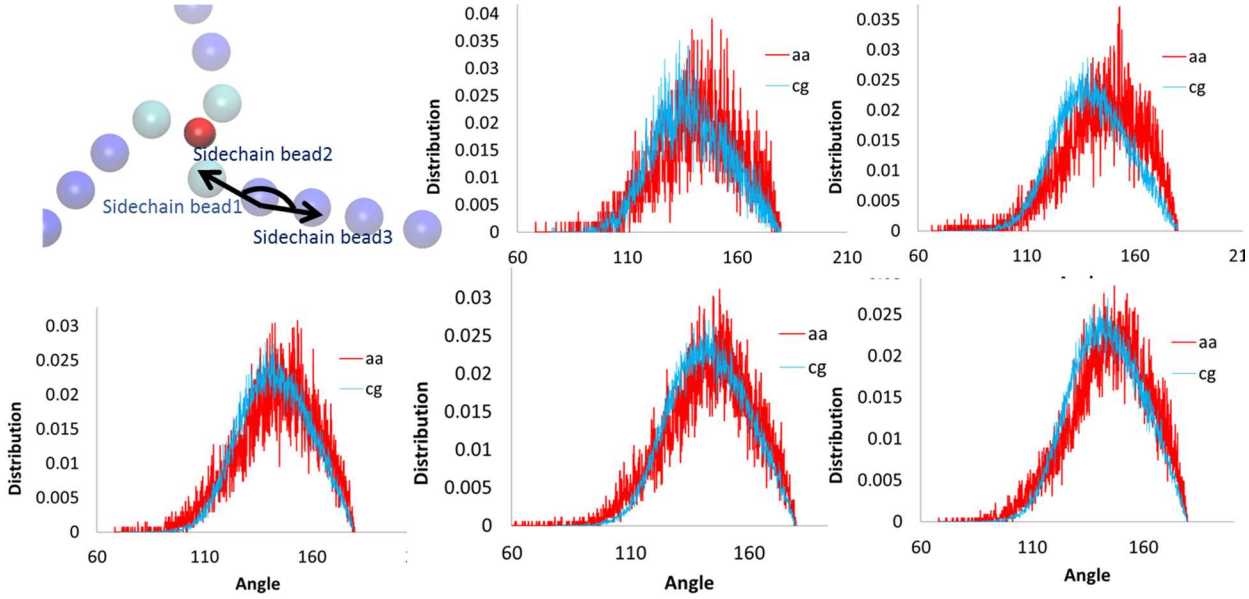


Figure 52. Comparison of AA and CG the first sidechain-sidechain angle distributions for PSLG (repeat unit = 3, 6, 9, 12, 15 displayed in the order of upper middle, upper right, bottom left, bottom middle and bottom right). The upper left subfigure elucidates that the first sidechain-sidechain angle is the space between the sidechain-sidechain bonds (showing in black arrows). In the rest subfigures, red line is the obtained from AA simulation; blue line is obtained from CG simulation.

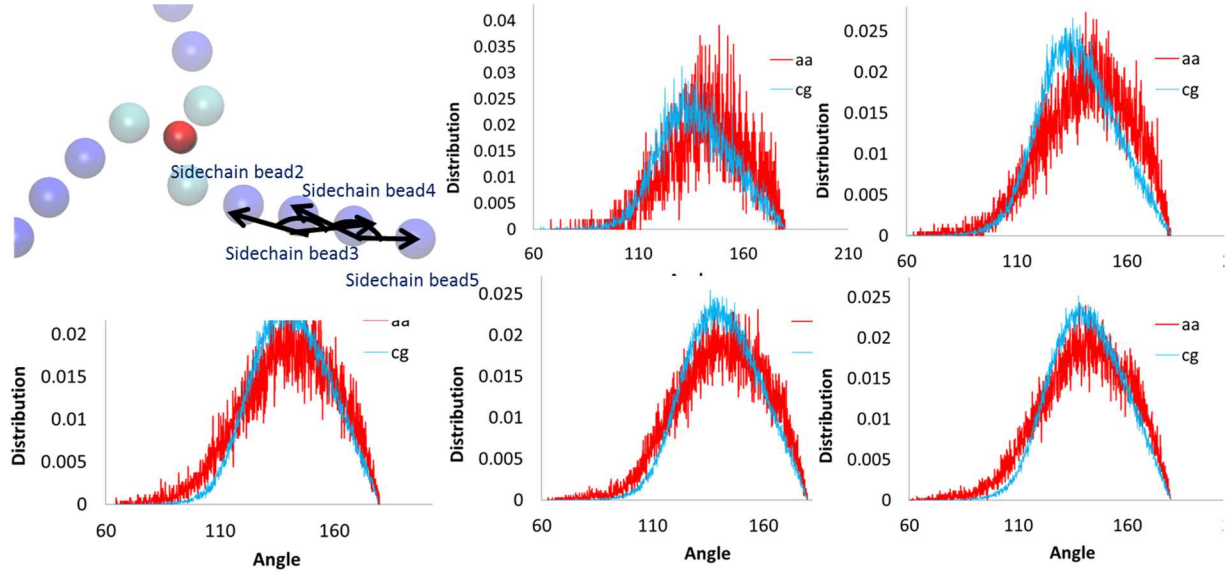


Figure 53. Comparison of AA and CG the second sidechain-sidechain angle distributions for PSLG (repeat unit = 3, 6, 9, 12, 15 displayed in the order of upper middle, upper right, bottom left, bottom middle and bottom right). The upper left subfigure elucidates that the second sidechain-sidechain angle is the space between the sidechain-sidechain bonds (showing in black arrows). In the rest subfigures, red line is the obtained from AA simulation; blue line is obtained from CG simulation.

We did notice that in several comparison cases of the AA and CG bonded distribution the deviation looked significant. For example, those results shown in Figure 51 for the case of 9, 12 and 15 repeat units. Nevertheless, the parameter tuning to make the difference smaller would usually change other properties or distributions. The eventual parameters were therefore a compromise and considered as overall optimal. Thus, though in some of the bonded distribution measurements the AA and CG results did not match exactly, overall our CG and AA models had similar bonded distributions and the model performance was considered acceptable.

In Figure 55 we report measurements of Rg of PSLG with repeated units from 3 to 15 when we placed the PSLG molecule in THF, although experimental values are not available at these sizes. In general the Rg values from our CG model was 15% larger than those from the AA simulations, which was considered acceptable.

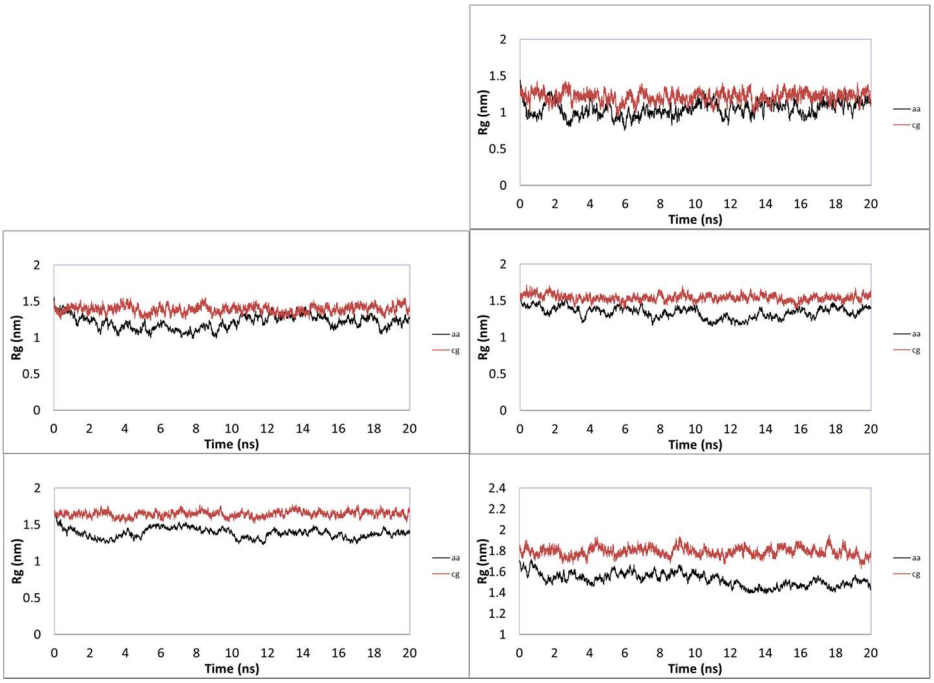


Figure 54. Radius of gyration of PSLG polymer in THF as a function of time. Repeat unit = 3, 6, 9, 12, 15 were displayed in the order of upper right, middle left, middle right, bottom left and bottom right.

Next, a CG model of a PSLG polymer molecule containing 240 repeat units was considered, which has a size comparable to the available experimental results [145, 146] where THF was used as solvent. We measured PSLG's size and shape, and compared to the experimental results when available. We measured Rg of PSLG as function of time during a 40 ns conventional MD simulation. As seen in Figure 55, the Rg value stabilizes after 10 ns of simulation time. The Rg had a mean value of 9.93 with standard deviation of 0.28. The observed minimum value of Rg was 9.30, and the observed maximum value of Rg was 10.57.

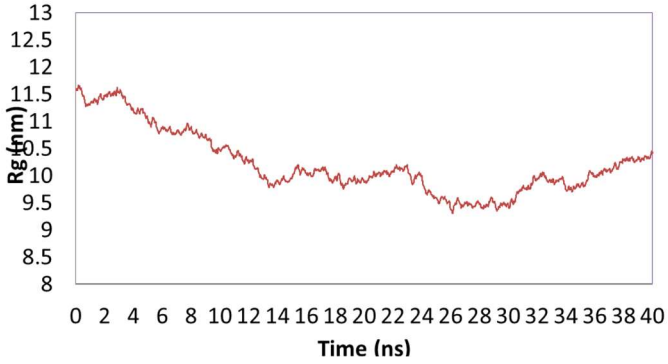


Figure 55. Radius of gyration of PSLG as a function of time. Repeat unit = 240

We also measured the shape of the PSLG polymer (Figure 56). It is possible to discuss average moments of inertia of the PSLG molecule over the last 30 ns simulation is 1:3.34:4.11,

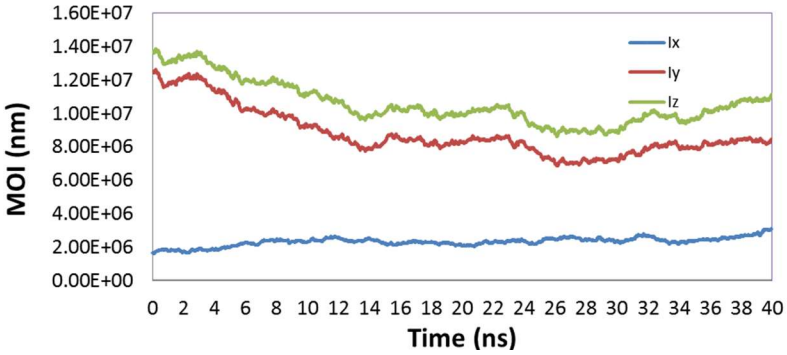


Figure 56. Moment of inertia of PSLG as a function of time. Repeat unit = 240

indicating a prolate ellipsoidal shape. In other words, the PSLG molecule did not coiled in THF during the simulation and the overall geometry was kept as an elongated rod.

Lastly, we compared the size and shape of PSLG obtained from CG MD simulation to the experimental results provided by Poche *et al*. [145]. Table 11provides a summary of data comparison for PSLG from CG MD simulation and experiment. The weight-average molecular weight (Mw) and the Rg were determined by static light scattering (SLS). As each repeat unit of PSLG has a molecular weight M_0 = 382, the PSLG in the experiment had about 251 repeat units (96000/382 = 251), which was more than the number of PSLG repeat unit in simulation. For modeling convenience, the PSLG for MD simulation has 240 units. Therefore, the Rg from experiment was the result for PSLG with 251 repeat unit, which should be higher than the simulation value. The characteristic end-to-end length of PSLG could be obtained either from Mw or Rg: L_M = Mw*0.15 and $L_{Rg} = 12^{1/2}$Rg. Although using L_M is more reasonable in the experiment measurement since in experiment a weight-averaged length was compared to the weight-averaged mass, we were more interested in L_{Rg} as this value can be used to compare the CG simulation with experiment in the term of characteristic end-to-end length. The thermodynamic diameter d, which is another shape parameter, can also be calculated from either $A_2 = \pi N_A d L^2/(4M^2)$ or Rg, The L_{Rg}, dRg, both of which were derived from Rg, obtained from the simulation were smaller than the results of experiment and deviation was acceptable. Like experiment results, in the simulation, the obtained L_{Rg} was much greater than d_{Rg}, suggesting the modeled PSLG molecule was not coiled but had a long rod-like structure.

Table 11. Summary of data comparison for PSLG between CG MD simulation and experiment

	Molecular Weight	Rg (nm)	L_M	L_{Rg}	d	d_{Rg}
Martini-CG	91700	9.93 ± 0.28	36.0	34.4	4.44	4.87
Experiment*	96000	13 ± 2.0	37.7	45.0	4.5	3.1

* Experiment values are from [145]
The unit of L_M, L_{Rg}, d and d_{Rg} is nm

3.4.3 Production run

We considered CG systems of 6 PSLG molecules initially perpendicular to a decane/water interface coated with hydrophobins ("perpendicular system" in short) or parallel to such interface ("parallel system" in short). Representative snapshots of these systems are presented in Figure 57. After 1000 ns simulation, some of the PSLG molecules approached to and then attached to the EAS hydrophobin monolayer, but didn't penetrate it. After the attachment, the PSLG molecules didn't leave the monolayer again, but their binding orientations differed from the initial alignment in the perpendicular alignment system. Both perpendicular and parallel alignments showed similar tendencies, with the backbone of the PSLG parallel to the plane of the hydrophobin monolayer. As seen in the snapshots, there were more PSLG molecules attached in perpendicular system than the parallel, which may be because all of the 6 PSLG molecules were close to the hydrophobin monolayer compared to 3 in the parallel system.

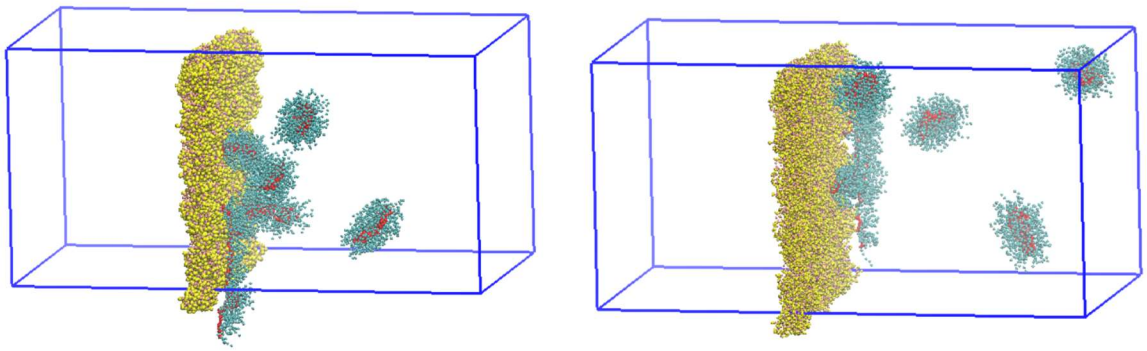


Figure 57. Snapshots of the final trajectory. Left: PSLG molecules initially perpendicular to the interface; right: PSLG molecules initially parallel to the interface. The molecule of PSLG is represented by red beads (for backbone) and cyan beads (for sidechains). The molecule of EAS hydrophobin is represented by pink beads (for backbone) and yellow beads (for sidechains). PSLG was solvated in decane on the right side. Water box was on the left side. For clarity, decane and water molecules were omitted.

In order to determine whether the molecules of interest (hydrophobin and PSLG) were at an overall equilibrating state, firstly we performed a Root Mean Square Deviation (RMSD)

analysis for the hydrophobin monolayer and the PSLG molecules to check the structural stability of the proteins. The analysis was carried on the system simulating two different concentrations of PSLG, start alignments of PSLG and in two different solvents. For example, Figure 58 displays the RMSD of the PSLG at low concentration solvated in decane. Hydrophobins in perpendicular system is in black line; hydrophobins in parallel system is in red line. PSLG in perpendicular system is in green line; PSLG in parallel system is in purple line. The RMSD calculation considered all of the 6 PSLG molecules as a whole. Some of the systems exhibit no significant change in the structures of the PSLG and hydrophobins after 400 ns, but other system required at least 800 ns of simulation time before reaching stable conditions.

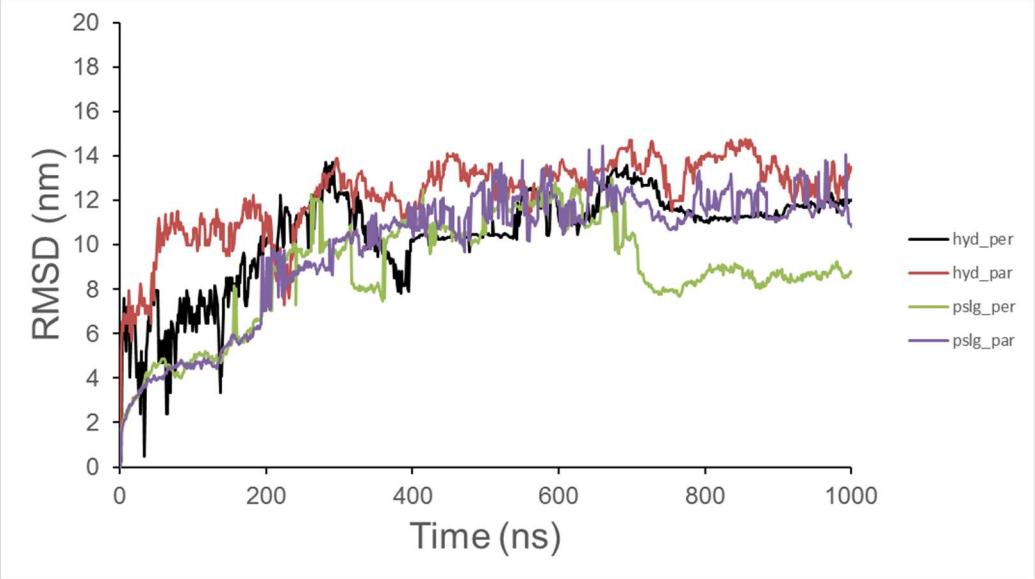


Figure 58. Evolution of the Root Mean Square Deviation (RMSD) of hydrophobin monolayer and PLSG considering different starting alignments for PLSG. Hydrophobins in perpendicular system is in black line; hydrophobins in parallel system is in red line. PSLG in perpendicular system is in green line; PSLG in parallel system is in purple line. Systems with 9 PSLG polymers and/or THF solvent showed the tendency, indicating the structure changes of PLSG for all tested systems were insignificant after several hundred nano-seconds.

Then we monitored radii of gyration of hydrophobin monolayer and PSLG for both perpendicular and parallel systems. On average the Rg values were stable during the 1000 ns

simulation in both perpendicular and parallel systems, displayed in black and red lines in Figure 59. The obtained values of Rg for the hydrophobin monolayer were close to each other in both systems, indicating that the size of hydrophobin monolayers was not affected by the alignment of PSLG. In the meantime, the average size of PSLG did not fluctuate significantly. Systems with high PSLG concentrations (9 PSLG polymers) solvated in decane or THF solvent showed similar trends.

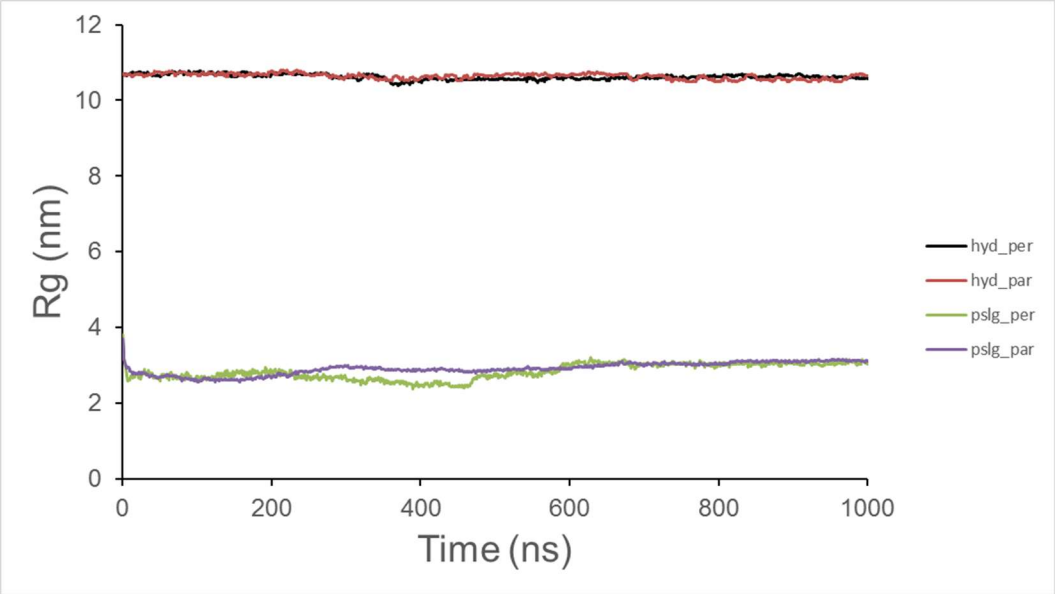


Figure 59. Radii of gyration of Hydrophobin monolayer and PSLG over the course of the 1000 ns simulation for considering different starting alignments for PLSG. Hydrophobins in perpendicular system is in black line; hydrophobins in parallel system is in red line. PSLG in perpendicular system is in green line; PSLG in parallel system is in purple line. Systems with 9 PSLG polymers and/or THF solvent showed the tendency, indicating the structure changes of PLSG for all tested systems were insignificant after several hundred nano-seconds.

The Rg of PSLG gives insights into the equilibrium of the whole hydrophobin monolayer, as well as the PSLG in average, but the size fluctuation is not the same for all molecules of the PSLG. To identify these differences within PSLG molecules, we examined the Rg values for individual PSLG chains. The upper part of Figure 60 reveals that when PSLG is solvated in decane, the Rg of each PSLG molecule differs from each other. There were two PSLG molecules stretched, size of which increased remarkably, after attaching to the

hydrophobin monolayer. Interestingly, other PSLG molecules attached to the hydrophobin monolayer did not show such a notable size increase. As both perpendicular and parallel systems had this phenomenon, the change of the size might be independent from the initial alignment situation. Limitations in cross-section area might be a reason why not all PSLG chains attach to the monolayer, and so some of the PSLG did not have enough space to stretch. Although the size of two PSLG increased, the increase was limited and became stable over time. All other PSLG molecules had a relatively stable size throughout the simulation. In contrast, in the case of THF as solvent, the Rg of PSLG molecules did not have significant change over time.

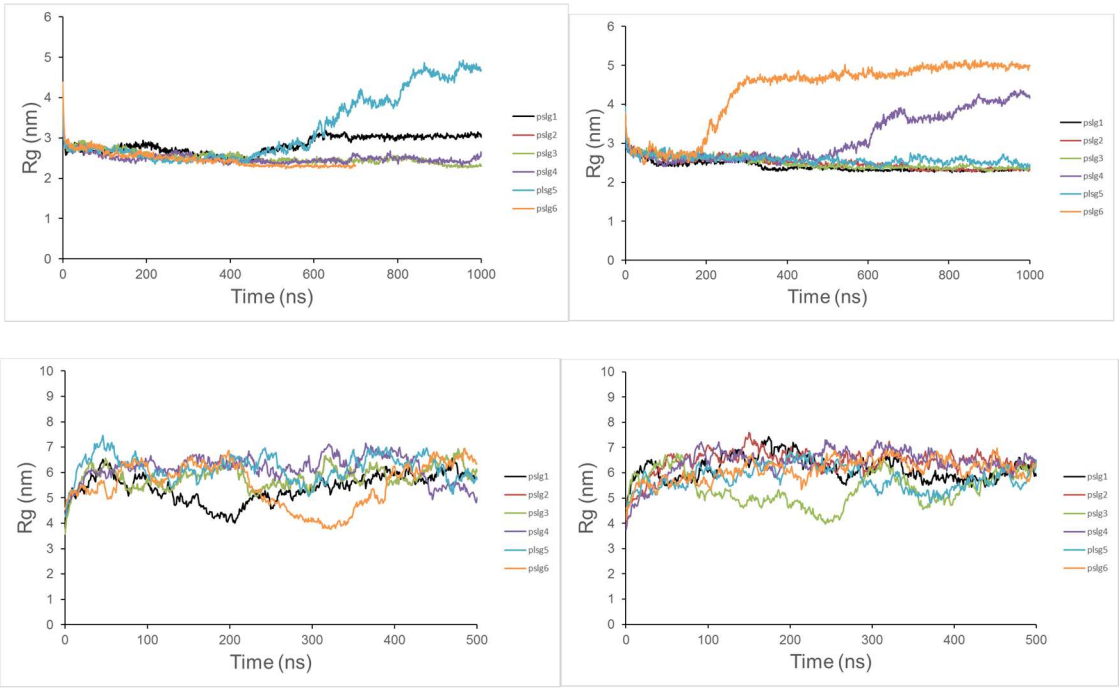


Figure 60. Radius of gyration of each PSLG molecule over the course of the 1000 ns simulation for considering different starting alignments for PLSG. Left: PSLG molecules initially perpendicular to the interface; right: PSLG molecules initially parallel to the interface. Upper: PSLG in decane; bottom: PSLG in THF.

We then probed the shape of the PSLG molecules for PSLG in perpendicular or parallel alignment at low or high concentration solvated in decane or THF. The shape of PSLG was characterized by moment of inertia (MOI). The MOI of each PSLG molecule where six PSLG

molecules were initially perpendicular or parallel to the interface are shown in Figure 61 and Figure 62. While 2 of the PSLG molecules in each system had an elongated shape, the other 4 approximately kept their initial shape. The PSLG molecules which had a remarkably shape change were the same ones that had a great size increase. Both perpendicular and parallel systems had this phenomenon, indicating the change of the size might be independent from the initial alignment situation. No matter if the shape changed significantly or not, the ratio of MOI implied each PSLG still had a prolate eclipse shape, though the prolate level of each molecule varied.

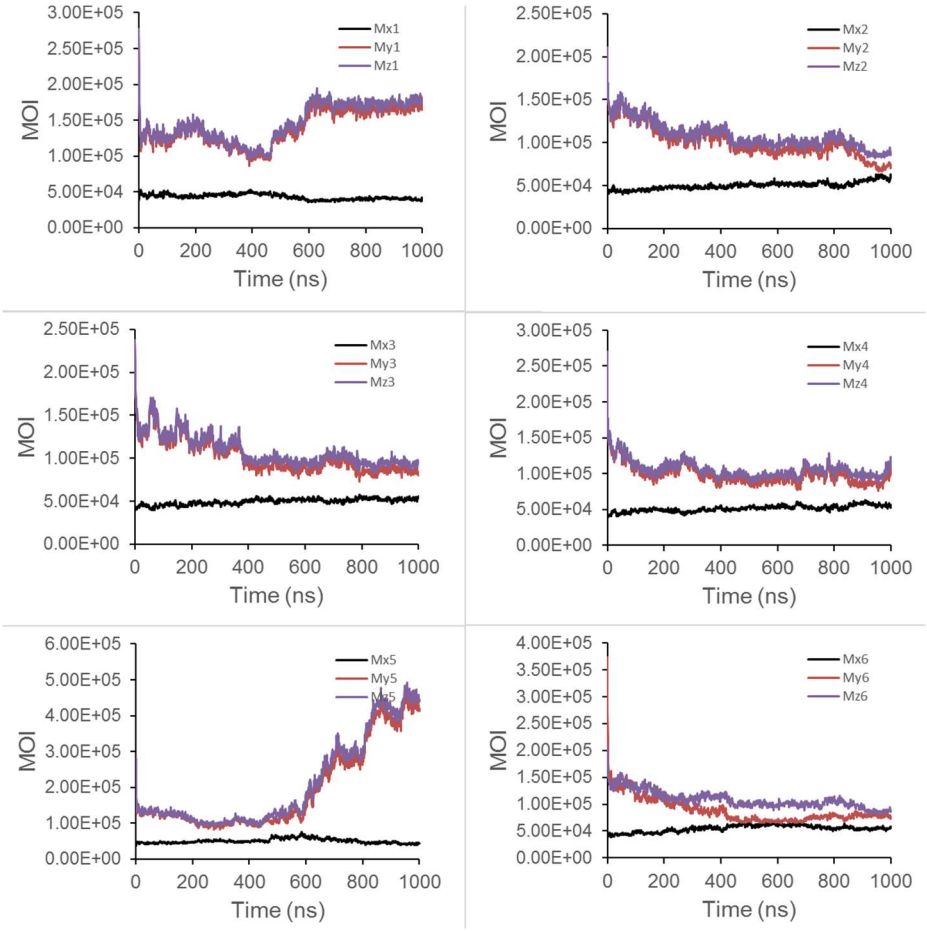


Figure 61. Moment of inertia (MOI) of each PSLG molecule where PSLG molecules were initially perpendicular to the interface. The components of MOI around the x (Mx), y (My) and z (Mz) axes are plotted.

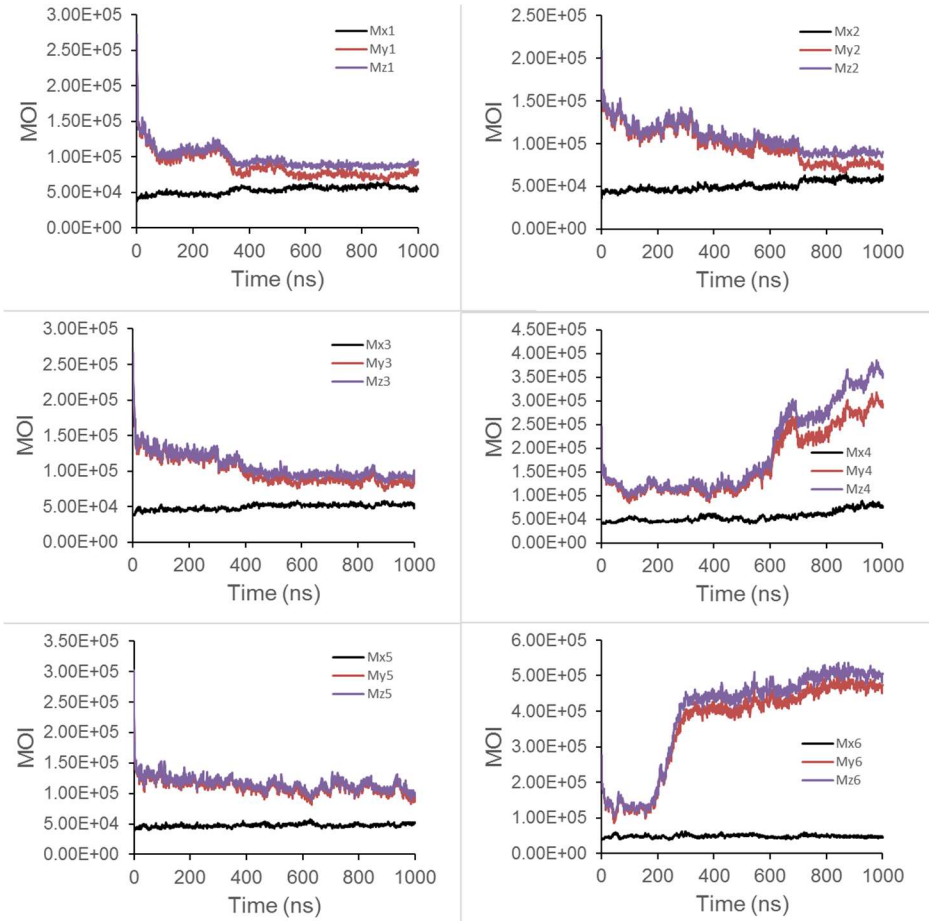


Figure 62. Moment of inertia (MOI) of each PSLG molecule where PSLG molecules were initially parallel to the interface. The components of MOI around the x (Mx), y (My) and z (Mz) axes are plotted.

When PSLG stays at or near the decane (THF)/water interface coated by a hydrophobin monolayer, the density profiles derived from the last 700ns in both perpendicular and parallel simulations are plotted in Figure 63. The curves of oil and water were sharp, while the hydrophobin peak was narrow and the width of the hydrophobin density distribution correlates to the Rg of the single hydrophobin molecule, suggesting a hydrophobin monolayer. The density profile of PSLG demonstrated that PSLG preferred to stay at the oil side of the interface. Hence, the distribution of PSLG density profile is consistent with the observation of the snapshots in Figure 57, from where most parts of the PSLG molecule were seen in the oil, while only minor parts were in the hydrophobin monolayer but never penetrated the layer. Moreover, as expected,

both the head and tail of PSLG were closer to the oil side of the oil/water interface and there was no significant difference between each other. Therefore the backbone chain of the PSLG were parallel to the plane of hydrophobin monolayer in either initially perpendicular or parallel systems.

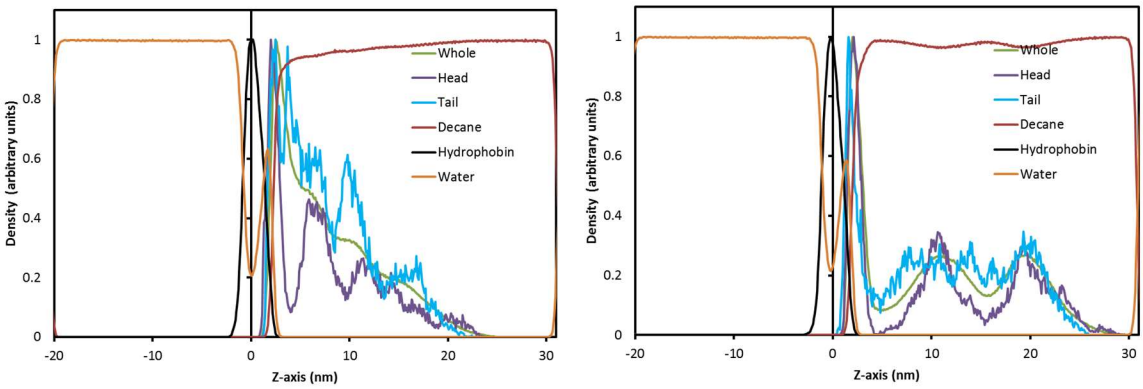


Figure 63. Density profiles for the PLSG molecules in hydrophobin coated oil/water interfaces. The profiles are for the whole EAS molecule (green), the head groups of the PSLG molecule (purple), and the tail groups of the PLSG molecule (light blue), decane (red), hydrophobin (black), water (orange). Left: PSLG molecules initially perpendicular to the interface; right: PSLG molecules initially parallel to the interface. The density profile of each species is normalized by dividing by the maximum value of their respective local density in the simulation box. The z coordinate was set to 0 at the maximum density of the hydrophobin monolayer.

Last, the probability distribution of the angle between PSLG backbone axis and the z-axis was probed to give an intuitive measurement of the PSLG alignment. The result, which excluded first 100ns simulation, was obtained from a series of simulations of the perpendicular or parallel system, low or high concentration, and in decane or THF. As shown in the left subfigure of Figure 64, the distribution plots reveal each initially perpendicular low-concentration PSLG in decane, as well as the averaged distribution. Although one PSLG molecules was perpendicular to the z-axis, the averaged PSLG backbones were evenly distributed. In the right subfigure, the distribution plots reveal each initially parallel low-concentration PSLG in decane, as well as the averaged distribution. The distribution shows the PSLG molecules were overall slightly

perpendicular to z-axis (parallel to the hydrophobin monolayer), but this is mainly because of one PSLG molecule (shown in the orange dots).

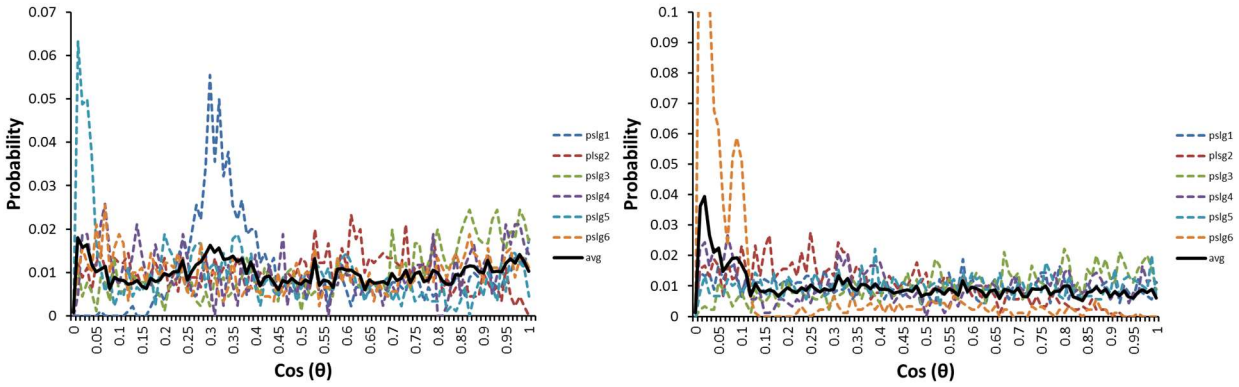


Figure 64. Probability distribution of the angle between PSLG backbone axis and the z-axis. The distribution plots reveal each initially perpendicular (left subfigure) or parallel (right subfigure) low-concentration (6 PSLG molecules in the simulation box) PSLG in decane. The black solid line is the average distribution of the total 6 PSLG molecules over the whole simulation time excludes the first 100 ns.

Unlike the results in decane, the angle distribution of the low-concentration PSLG in THF showed a certain relation to their initial alignments. Figure 65 shows in THF the probability distribution of the angle between PSLG backbone axis and the z-axis. As seen in the left subfigure of Figure 65, the PSLG backbones were mainly parallel to z-axis (perpendicular to the

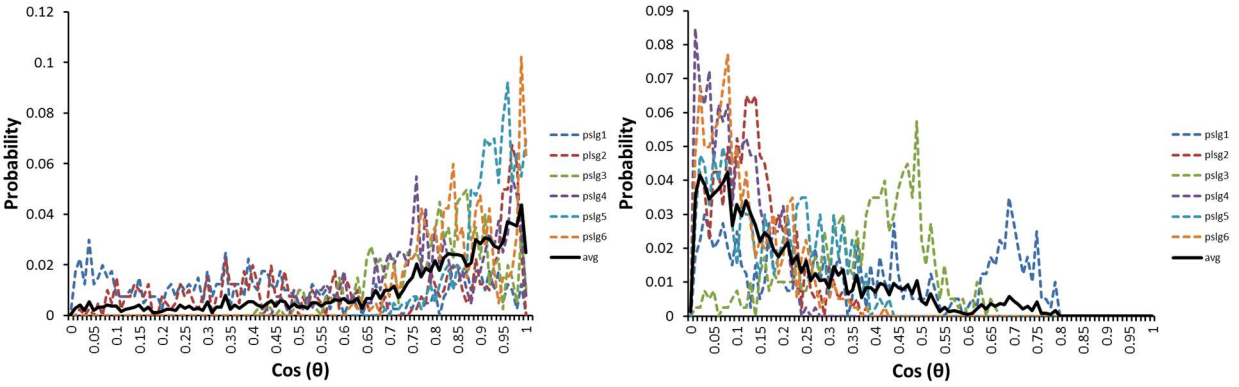


Figure 65. Probability distribution of the angle between PSLG backbone axis and the z-axis. The distribution plots reveal each initially perpendicular (left subfigure) or parallel (right subfigure) low-concentration (6 PSLG molecules in the box) PSLG in THF. The black solid line is the average distribution of the total 6 PSLG molecules over the whole simulation time excludes the first 100 ns.

plane of hydrophobin monolayers), whereas in the right subfigure of Figure 65, the PSLG backbones were mainly perpendicular to z-axis (parallel to the plane of hydrophobin monolayers). The distribution results for both cases were correlated to the initial alignment of the PSLG, compared to the cases in decane, indicating the hydrophobicity of the solvent might be able to affect the final status of the PSLG molecules.

The distribution plots in Figure 66 reveals angle distribution of high-concentration PSLG in decane. For the initially perpendicular alignment, the average distribution of high-concentration PSLG in decane is evenly distributed in general, which is the same as low-concentration. But for the initially parallel alignment, as displayed in the right subfigure of Figure 66, about have of the PSLG molecules were perpendicular to z-axis (parallel to the plane of hydrophobin monolayers) in general, making the overall alignments of PSLG perpendicular to z-axis.

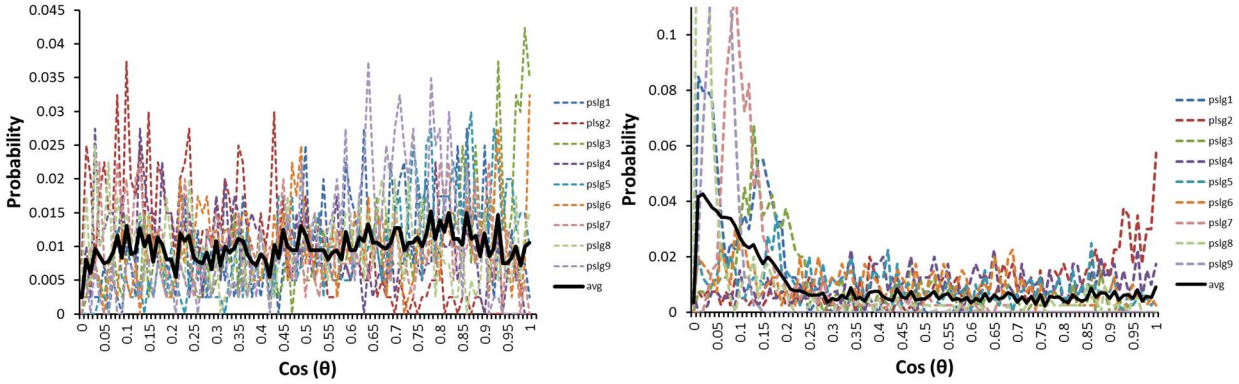


Figure 66. Probability distribution of the angle between PSLG backbone axis and the z-axis. The distribution plots reveal each initially perpendicular (left subfigure) or parallel (right subfigure) high-concentration (9 PSLG molecules in the box) PSLG in decane. The black solid line is the average distribution of the total 9 PSLG molecules over the whole simulation time excludes the first 100 ns.

The distribution plots in Figure 67 reveals angle distribution of high-concentration PSLG in THF. As seen in the top subfigure of Figure 67, the PSLG backbones initially perpendicular to the hydrophobin monolayers were mainly parallel to z-axis (perpendicular to the plane of

hydrophobin monolayers), which is the same as low-concentration PSLG in THF. However, in the initially parallel case, as seen in the bottom subfigure of Figure 67, high-concentration PSLG molecules went into two extremes: either perpendicular or parallel to the z-axis, making the overall distribution even. Compared to the result of initially parallel low-concentration PSLG in THF, the "extra" PSLG molecules tend to be parallel to z-axis (perpendicular to the plane of hydrophobin monolayers).

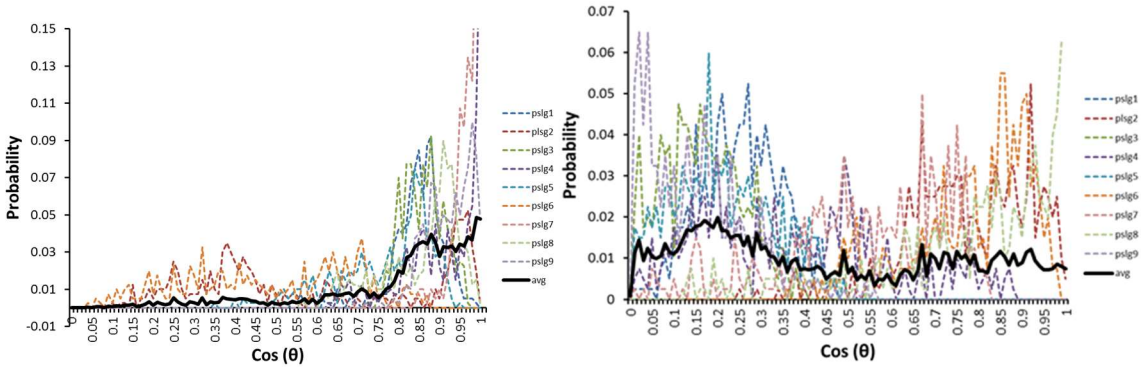


Figure 67. Probability distribution of the angle between PSLG backbone axis and the z-axis. The distribution plots reveal each initially perpendicular (top subfigure) or parallel (bottom subfigure) high-concentration (9 PSLG molecules in the box) PSLG in THF. The black solid line is the average distribution of the total 9 PSLG molecules over the whole simulation time excludes the first 100 ns.

3.5 Coarse-grained parameterization of P3HT

Poly(3-hexylthiophene), P3HT is a commercially available semiconducting polymer often used as a model system, although trails the helical polypeptide PSLG in terms of rigidity, monodispersity, and solubility. MD simulations with CG models for the systems of interest investigation can reach to larger dimensions for longer time scales than AA models. In order to validate the CG model, comparing CG results with experimental and AA results is necessary. Therefore we first built an OPLS-AA model of P3HT, of which parameters were provided by different works [84, 160-162]. But the provided OPLS-AA parameters of P3HT were all for the

simulation of P3HT crystals or membrane attached to solid interface. To obtain the correct solution property, the AA model of P3HT for the solid state needs to be further verified.

Then similarly to the PSLG modeling, based on the AA trajectory, several strategies of group coarse-graining had been proposed and carefully examined. For simplicity the selection process is not included in this work. The tentative coarse-graining plan was shown in Figure 68. The strategy to model rings is to include as three CG beads as necessary in order to keep the ring geometry. With this more detailed mapping, enough geometrical detail is kept to mimic the geometry of central repeat unit. Two C1 beads are assigned to represent alkane tails in the red circle; two C4 beads are assigned to represent carbon rings in blue; one N0 bead is assigned to represent sulfur in green. The ring beads is in a special type provided by the Martini to allow effective interaction size and strength for ring-ring interactions, which is reduced compared to the normal set. The ring geometry will be kept by constraints, while other structures are kept by bond and angle potentials.

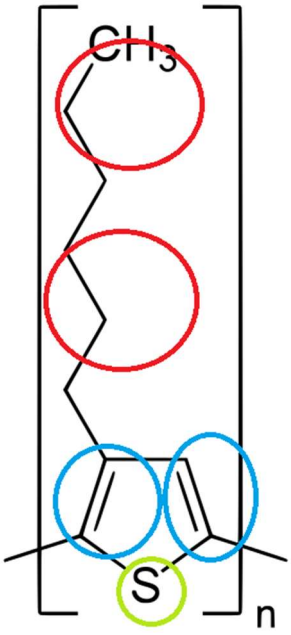


Figure 68. A simpler representation may be possible; however, the current proposal allows for stable simulations with an integration time step up to 20 fs (actual simulation time).

3.6 Research progress conclusion

We have built AA and CG models of PSLG based on what we learned about the PSLG. OPLS-AA force field and Martini force field were applied to build the AA and CG model respectively. The necessary bonded and nonbonded parameters required to define the PSLG repeat unit were proposed according to the standard OPLS-AA and an optimization work for the long carbohydrate chain. As a Martini CG model of THF was not available, we also built a CG model for this molecule.

Conventional MD simulations were then carried out to validate the CG model of THF as well as the AA and CG models of PSLG, and then for the formal production run. The thermodynamic integration was carried out to test the free energies of the THF model. The conventional MD simulation of PSLG polymers with small number of repeat unit in THF was performed first to validate the bonded distributions, followed by a simulation of a PSLG polymer having 240 repeat units in a large THF-solvated box to further improve the CG model, and lastly with the "total verified" CG model of PSLG, more complicated system including the interface between multiple PSLG molecules and the hydrophobins encapsulating them was created to test alignment and packing of semiflexible polymers by hydrophobins.

PSLG model having size comparable to the experiment have been created and tested. We concluded that the obtained results from CG simulation were in accordance with to the experimental results. Characteristic values of size and shape obtained from CG models were also qualitatively captured the experimental observations of the size and shape of the PSLG solvated in THF.

In the 500-1000 ns production runs of the perpendicular and parallel systems, low or high concentration of PSLG in decane or THF, we probed the alignments of PSLG molecules near a

hydrophobin coated hydrophilic-hydrophobic interface by measuring the angle distribution between PSLG backbone axis and the z-axis, which is the vector of the plane of EAS monolayer. The initial alignment (perpendicular or parallel to the z-axis), PSLG concentration (various number of PSLG in the simulation box) and solvent type (decane as more hydrophobic solvent, THF as less hydrophobic solvent) are chosen as the impact factors that all affect the alignment of PSLG during the 500 – 1000 ns simulation. PSLG at low concentrations in THF keep their original alignment, but PSLG at high concentrations in THF have a somewhat different behavior. In contrast, PSLG at low concentrations in n-decane lose their original alignment, whereas PSLG at high concentrations in n-decane also lose their original alignment. The effects of these three impact factors on the PSLG alignment are correlated and should not be considered alone.

Chapter 4. Conclusions and Recommendations for Future Studies

4.1 Conclusions

We at first investigated the properties of EAS, a class I hydrophobin, near gas/water and oil/water interfaces. Here we report results of the interfacial properties of these systems, as determined from classical molecular dynamics (MD) simulations and potential of mean force (PMF) calculations using both all-atom (AA) and coarse-grained (CG) representations of the different species. Relevant properties, such as free energies, density profiles, structure of hydrophobins were probed and discussed in this study. Because of the high structural and functional similarity of hydrophobins, remarkably, most research progress on EAS is transferable to other hydrophobins.

According to the PMF calculation results, the EAS molecule is likely to stay at the water side of the interfaces as indicated by the presence of deep free energy minima. When the value of PMF is arbitrarily set to 0 $kJ \cdot mol^{-1}$ in water, the free energy minimum obtained from atomistic simulation is -28.9 $kJ \cdot mol^{-1}$, while the minimum obtained from CG simulation is -25.0 $kJ \cdot mol^{-1}$. These two values, along with the shape of the curves are close to one another, indicating that the PMF results from the Martini force field are analogous to those from OPLS-AA force field. Between the free energy minimum and PMF in the gas phase far from the interface, the observed gap is large (>1000 $kJ \bullet mol^{-1}$), inferring that hydrophobins prefer not to stay in the gas phase.

Conventional MD simulations also confirm that EAS prefers to stay at the water side of the interfaces. Our snapshots and density profile measurement of all component in the simulation box showed that once the EAS was adsorbed to the interfaces it did not leave interfaces. This is consistent with the thermodynamical trends we observed from PMF calculations. This preference implies the good capability of hydrophobin to stabilize the hydrophobic/hydrophilic interface as

a bio-surfactant. Our PMF calculations also suggest that the adsorption behavior of EAS at the interface is strong and irreversible.

After EAS bound to the interfaces, the simulated radius of gyration (Rg) was close to the experimental value and experience small variations. The overall structure of EAS became more stable at the interface than in the bulk water. In addition, the secondary structure change of EAS was consistent with the experimental observations. We measured the average Rg of single EAS molecules and the interfacial tension in each simulation box, and found that while the average size of EAS did not change significantly as the number increased, the interfacial tension decreases as the concentration of EAS at the interface increases. Finally, we performed MD simulations using Martini CG models to gain insight into the stability of nm-sized 'blobs' formed by the assembly of hydrophobin around small oil blobs. The class I hydrophobin EAS molecules were initially placed around a pre-equilibrated benzene blob with *cylindrical* shape in bulk water, then the MD simulation was ran 500 -2000ns. Beginning from the tiny benzene blob that had 0 benzene and 4 EAS molecules (strictly speaking it was not a benzene blob as the model did not contain benzene), we then made the blob model larger and larger in size, and finally built a large blob (3334 benzenes, 109 EAS for single layer simulation or 218 EAS for double layer simulation) with a size (~10 nm in Rg) that was comparable to the smallest blob determined by experiment. The hydrophobin molecules were successfully assembled at the benzene blob surface, and the blob size was stable over the course of simulation. The blob shape was determined to be slightly elongated through the moment of inertia measurement (for small blobs) or asphericity characterized by the gyration tensor (for large blobs). The density distribution profile, which is the density as a function of the distance from the center of mass of the blob, showed that benzene molecules stayed inside the center of blob, coated by the hydrophobic parts

of the EAS, and then hydrophilic parts. Lastly water was at the outside of the blob and cannot penetrated into it.

In our second project, we built AA and CG models of PSLG using the OPLS-AA and Martini force fields. The necessary bonded and nonbonded parameters required to define the PSLG repeat unit were proposed and written into the residue database for the topology building, according to the standard OPLS-AA and an optimization work for the long carbohydrate chain. We also developed a CG model for THF with Martini force field. During the CG modeling of THF and PSLG, we tried to assign various bead types to represent functional groups of the THF and PSLG. Then the bonded interaction parameters were written to finish the topology building. PSLG polymers with repeat units ranging from 3 to 240 were created respectively to optimize the model of PSLG with conventional MD simulations and thermodynamic integration. Conventional MD simulation was carried out to validate the CG model of THF as well as the AA and CG models of PSLG, and then for the later production run. Thermodynamic integration was carried out to test the free energies of the THF model. The conventional MD simulation of PSLG polymers with small number of repeat unit in THF was performed first to validate the bonded distributions, followed by a simulation of a PSLG polymer having 240 repeat units in a large THF-solvated box to further improve the CG model, and lastly with the "total verified" CG model of PSLG, more complicated system including the interface between multiple PSLG molecules and the hydrophobins encapsulating them was created to test alignment and packing of semiflexible polymers by hydrophobins. The PSLG molecules were solvated in decane in a periodic simulation box, where the backbone chain orientation of all PSLG was either perpendicular or parallel to the hydrophobin monolayer at the oil/water interface.

The CG model of THF was examined first. After substantially tweaking the parameters, the simulated free energies of evaporation from AA and CG results became very close to each other, but had differences from the calculated experiment results. However, the simulated free energies of hydration from AA and CG results were in very good agreement with experiment results. The obtained bulk densities from both AA and CG were in excellent agreement with experiment as well. CG parameters for the bond and angle potentials were optimized by comparing bonded distributions with AA simulations of THF mapping them to the CG representation. Results from CG simulations were in accordance with AA results.

The AA and CG models of PSLG were then validated. The bonded and nonbonded parameters for the PSLG polymer were optimized by comparing CG and AA results for properties such as RMSD, bond length and angle distributions. Three kinds of bond length and four kinds of angle distributions for each PSLG polymers having repeat units from 3 to 15 were probed. We concluded that the obtained results from CG simulation were in agreement with AA results. In particular, the angle distribution situation was in agreement with the experimental observation. Then in the CG simulation of PSLG with 240 repeat units, the Rg of PSLG molecule was reasonably close to the reported experimental value. Other characteristic values of size and shape obtained from CG models were also in qualitatively agreement with experimental observations of the size and shape of PSLG solvated in THF.

In a series of 500-1000 ns production runs, the effects of the initial alignment (perpendicular or parallel to the z-axis), PSLG concentration (various number of PSLG in the simulation box) and solvent type (decane as more hydrophobic solvent, THF as less hydrophobic solvent) on the PSLG alignment were tested. Some PSLG molecules were observed approaching to the EAS hydrophobin monolayer and then attaching to it, but the attached PSLG molecules

did not penetrate the hydrophobin monolayer. Two PSLG molecules attached to the hydrophobin monolayer and stretched, causing the size of these two PSLG molecules to increase remarkably; however other PSLG molecules that attached to the hydrophobin monolayer did not show such a notable size increase. While two of the PSLG molecules in each system had an elongated shape, the other four kept their initial shape. Both perpendicular and parallel systems had this phenomenon, indicating the change of the size might be independent from the initial alignment situation. No matter if the shape changed significantly or not, the ratio of MOI implied each PSLG still had a prolate eclipse shape, though the prolate level varied. The results indicate that all of the three impact factors: initial alignment, PSLG concentration and solvent type may affect the alignment of the PSLG at the EAS monolayer. The effects of these three impact factors on the PSLG alignment are correlated and should not be considered alone.

4.2 Recommendations for future studies

Our current MD simulation result has suggested that hydrophobins could be used as alternatives to normal surfactants in a number of applications. Further work is required to optimize the current hydrophobin EAS and semiflexible polymer PLSG model built with Martini FF. Also, new models for class II hydrophobins and for a P3HT chain should be developed respectively to expand our simulation work. If time and/or length of the simulation need to reach to a significant large scale that even Martini CG method cannot access, a more CG method should be developed rather than the standard four-to-one mapping used in Martini models.

Martini FF will be applied to build class II hydrophobin HFBII model. Class II hydrophobins are able to form air bubbles (oil blobs) in a large variety of sizes ranging from nm to mm in Rg, whereas the aggregates formed by class I hydrophobins can only reach a nm-scale size. Hence, modeling a class II hydrophobin is thus more amenable for direct comparisons

between simulations and experiments. Replacing class I EAS with class II HFBII, we will perform similar MD simulations introduced in Section 2.2, using CG models to gain insight into the stability of the 'blobs' formed by the assembly of hydrophobin around small oil droplets. If the simulated result with HFBII model still shows that the blob has an ellipsoidal shape, consisting with the experiment observation of *cylindrical* blobs, our comparison will make more sense because of the hydrophobin class consistency. The PMF calculation in Chapter 2 will be repeated, but now for HFBII instead of EAS. Then the results of HFBII will be compared to EAS. The comparison will hopefully give us some insights on what is different between HFBII (which can make large *cylindrical* structures from nm to mm-scale) and EAS (which cannot make *cylindrical* structures greater than nm-scale).

The AA and CG modeling and characterization of P3HT introduced in Section 3.5 is still in progress. A tentative coarse-graining plan has been proposed by assigning bead types and bonded parameters to the functional groups. Next, the validation process of the P3HT model will almost be the same as PSLG. A Martini CG P3HT molecule with 10 repeat units will be created and placed in organic solvent for the bonded parameter distribution test, results of which will be compared to the AA results. After that test, some basic properties of the P3HT model such as Rg will be probed, and then we will perform a simulation of a P3HT polymer having >100 repeat units solvated in a large organic solvent box. Lastly, we will perform CG simulations of P3HT in organic solvent (either decane or THF), forming an interface with hydrophobins in water. The PSLG molecules will be solvated in decane in a periodic simulation box, where the backbone chain orientation of all PSLG will be either perpendicular or parallel to the hydrophobin monolayer at the oil/water interface.

At last, it is possible that we might need to consider time and length scales that are significantly larger than those accessible with Martini FF. In this case, Dissipative particle dynamics (DPD) simulations [163-165] using the LAMMPS [166] software will be conducted for PSLG/P3HT, the hydrophobins, water and the organic solvents. In this approach, several water molecules will be grouped and modeled as a single particle (a 'water' bead). Likewise, several molecules of organic solvent or gas will be modeled as one bead (a 'oil' bead or a 'gas' bead). All beads are the same size. A hydrophobin molecule will be modeled as a hard core on which hydrophobic and hydrophilic beads will be attached via FENE potentials [167]. Figure 69 shows a molecule of CU as well as the different particles to be considered in the DPD simulations. All the particles in these representations interact via softer-than-Lennard-Jones two-body interactions. In addition to the usual conservative

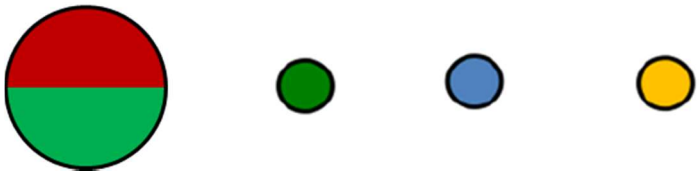


Figure 69. Schematic representation of particles to be considered in the DPD simulations. From left to right: CU molecule, consisting of a hard core (white) with hydrophobic (green) and hydrophilic (red) beads; oil bead (green); water bead (blue); and gas bead (yellow).

particle interactions, in DPD simulations dissipative and random forces also act. This is because hydrodynamic effects become important, as the DPD simulations can reach time scales on the order of microseconds. System sizes larger than those accessible to CG MD simulations can be studied via DPD simulations at the cost of losing molecular-level details. Simulations will be carried out in the NVT ensemble. As a first approximation, the parameters for the water beads, oil beads, hydrophilic and hydrophobic beads will be directly adopted from the DPD simulations

of Rekvig *et al*. [165] As the parameters used in this initial set of DPD simulations will be adopted from previous studies for slightly different systems, results from these simulations may only agree qualitatively (as opposed to quantitatively) with experiments. More accurate parameters for the water beads, oil beads, gas beads, and hydrophilic and hydrophobic beads can be determined by fitting these DPD parameters to reproduce properties obtained from all-atom MD simulations of similar systems (of much smaller sizes and for much shorter simulated times).

References

1. Lugones, L.G., Wosten, H.A.B., and Wessels, J.G.H., A hydrophobin (ABH3) specifically secreted by vegetatively growing hyphae of Agaricus bisporus (common white button mushroom). *Microbiology* 1998. 144(8): p. 2345-2353.

2. Wessels, J.G.H., Hydrophobins: Proteins that Change the Nature of the Fungal Surface, in *Advances in Microbial Physiology*, R.K. Poole, Editor. 1996, Academic Press. p. 1-45.

3. Wosten, H.A., Hydrophobins Multipurpose proteins. *Annual Review of Microbiology* 2001. 55: p. 625-646.

4. Russo, P.S., Blum, F.D., Ipsen, J.D., Abul-Hajj, Y.J., and Miller, W.G., The Solubility and Surface-Activity of the Ceratocystis-Ulmi Toxin Cerato-Ulmin. *Physiological Plant Pathology* 1981. 19: p. 113-126.

5. Russo, P.S., Blum, F.D., Ipsen, J.D., Abul-Hajj, Y.J., and Miller, W.G., A Novel Delivery Mechanism for Cerato-Ulmin in Elm Trees. *Proceedings of the Dutch Elm Disease Symposium and Workshop* 1981: p. 203-214.

6. Russo, P.S., Blum, F.D., Ipsen, J.D., Abul-Hajj, Y.J., and Miller, W.G., The Surface-Activity of the Phytotoxin Cerato-Ulmin. *Canadian Journal of Botany-Revue Canadienne de Botanique* 1982. 60: p. 1414-1422.

7. Brasier, C.M. and Gibbs, J.N., Origin of the Dutch Elm Disease Epidemic in Britain. *Nature* 1973. 242: p. 607-609.

8. Stevenson, K.J., Slater, J.A., and Takai, S., Cerato-ulmin - a wilting toxin of dutch elm disease fungus. *Phytochemistry* 1979. 18: p. 235-238.

9. Takai, S., Pathogenicity and cerato-ulmin production in ceratocystis ulmi. *Nature* 1974. 252: p. 124-127.

10. Schuren, F.H.J. and Wessels, J.G.H., Two genes specifically expressed in fruiting dikaryons of Schizophyllum commune: homologies with a gene not regulated by mating-type genes. *Gene* 1990. 90(2): p. 199-205.

11. Wosten, H.A., Schuren, F.H.J., and Wessels, J.G., Interfacial self-assembly of a hydrophobin into an amphipathic protein. *The EMBO Journal* 1994. 13(24): p. 5848-5854.

12. Wessels, J.G., Developmental Regulation of Fungal Cell Wall Formation. *Annual Review of Phytopathology* 1994. 32: p. 413-437.

13. Linder, M.B., Hydrophobins: Proteins that self assemble at interfaces. *Current Opinion in Colloid & Interface Science* 2009. 14(5): p. 356-363.

14. Sunde, M., Kwan, A.H., Templeton, M.D., Beever, R.E., and Mackay, J.P., Structural analysis of hydrophobins. *Micron* 2008. 39(7): p. 773-784.

15. Stringer, M.A. and Timberlake, W.E., A Toxin Involved in Dutch Elm Disease Is A Fungal Hydrophobin. *Plant Cell* 1993. 5: p. 145-146.

16. Stringer, M.A., Dean, R.A., Sewall, T.C., and Timberlake, W.E., Rodletless, a new Aspergillus developmental mutant induced by directed gene inactivation. *Genes & Development* 1991. 5(7): p. 1161-1171.

17. Bowden, C.G., Hintz, W.E., Jeng, R., Hubbes, M., and Horgen, P.A., Isolation and characterization of the cerato-ulmin toxin gene of the Dutch elm disease pathogen, Ophiostoma ulmi. *Current Genetics* 1994. 25(4): p. 323-329.

18. Lodish, H., Berk, A., Matsudaira, P., Kaiser, C., Krieger, M., Scott, M., Zipurksy, S., and Darnell, J., Molecular Cell Biology. 5th ed. 2004, New York: WH Freeman and Company.

19. Santos, C. and Labarère, J., Aa-Pri2, a single-copy gene from Agrocybe aegerita, specifically expressed during fruiting initiation, encodes a hydrophobin with a leucine-zipper domain. *Current Genetics* 1999. 35(5): p. 564-570.

20. de Vocht, M.L., Scholtmeijer, K., van der Vegte, E.W., de Vries, O.M.H., Sonveaux, N., Wosten, H.A.B., Ruysschaert, J.-M., Hadziioannou, G., Wessels, J.G.H., and Robillard, G.T., Structural characterization of the hydrophobin SC3, as a monomer and after self-assembly at hydrophobic/hydrophilic interfaces. *Biophysical Journal* 1998. 74: p. 2059-2068.

21. Lauter, F.R., Russo, V.E., and Yanofsky, C., Developmental and light regulation of eas, the structural gene for the rodlet protein of Neurospora. *Genes & Development* 1992. 6(12a): p. 2373-2381.

22. Spanu, P.D., HCF-1, a hydrophobin from the tomato pathogen Cladosporium fulvum. *Gene* 1997. 193: p. 89-96.

23. Talbot, N., Ebbole, D., and Hamer, J., Identification and characterization of MPG1, a gene involved in pathogenicity from the rice blast fungus Magnaporthe grisea. *Plant Cell* 1993. 5(11): p. 16.

24. Nakari-Setälä, T., Aro, N., Kalkkinen, N., Alatalo, E., and Penttilä, M., Genetic and Biochemical Characterization of the Trichoderma Reesei Hydrophobin HFBI. *European Journal of Biochemistry* 1996. 235(1-2): p. 248-255.

25. Muñoz, G., Nakari-Setälä, T., Agosin, E., and Penttilä, M., Hydrophobin gene srh1, expressed during sporulation of the biocontrol agent Trichoderma harzianum. *Current Genetics* 1997. 32(3): p. 225-230.

26. Zhang, L., Villalon, D., Sun, Y., Kazmierczak, P., and van Alfen, N.K., Virus-associated down-regulation of the gene encoding cryparin, an abundant cell-surface protein from the chestnut blight fungus, Cryphonectria parasitica. *Gene* 1994. 139(1): p. 59-64.

27. Nielsen, P.S., Clark, A.J., Oliver, R.P., Huber, M., and Spanu, P.D., HCf-6, a novel class II hydrophobin from Cladosporium fulvum. *Microbiological Research* 2001. 156(1): p. 59-63.

28. Bolyard, M.J. and Sticklen, M.B., Expression of a modified Dutch elm disease toxin in Escherichia coli. *Molecular Plant-Microbe Interactions* 1992. 5(6): p. 520-524.

29. St. Leger, R.J., Staples, R.C., and Roberts, D.W., Cloning and regulatory analysis of starvation-stress gene, ssgA, encoding a hydrophobin-like protein from the entomopathogenic fungus, Metarhizium anisopliae. *Gene* 1992. 120(1): p. 119-124.

30. Kwan, A.H., Winefield, R.D., Sunde, M., Matthews, J.M., Haverkamp, R.G., Templeton, M.D., and Mackay, J.P., Structural basis for rodlet assembly in fungal hydrophobins. *Proceedings of the National Academy of Sciences of the United States of America* 2006. 103(10): p. 3621-3626.

31. Hakanpaa, J., Paananen, A., Askolin, S., Nakari-Setala, T., Parkkinen, T., Penttila, M., Linder, M.B., and Rouvinen, J., Atomic resolution structure of the HFBII hydrophobin, a self-assembling amphiphile. *The Journal of Biological Chemistry* 2004. 279(1): p. 534-539.

32. Russell, R.B. and Barton, G.J., Multiple protein sequence alignment from tertiary structure comparison: Assignment of global and residue confidence levels. *Proteins: Structure, Function, and Bioinformatics* 1992. 14(2): p. 309-323.

33. Humphrey, W., Dalke, A., and Schulten, K., VMD: Visual molecular dynamics. *Journal of Molecular Graphics* 1996. 14(1): p. 33-38.

34. Kabsch, W. and Sander, C., Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 1983. 22(12): p. 2577-2637.

35. Joosten, R.P., te Beek, T.A.H., Krieger, E., Hekkelman, M.L., Hooft, R.W.W., Schneider, R., Sander, C., and Vriend, G., A series of PDB related databases for everyday needs. *Nucleic Acids Research* 2011. 39(suppl 1): p. D411-D419.

36. Kwan, A.H., Macindoe, I., Vukasin, P.V., Morris, V.K., Kass, I., Gupte, R., Mark, A.E., Templeton, M.D., Mackay, J.P., and Sunde, M., The Cys3-Cys4 loop of the hydrophobin EAS is not required for rodlet formation and surface activity. *Journal of Molecular Biology* 2008. 382(3): p. 708-720.

37. Moore, D., Fungal Morphogenesis. 1998, Cambridge: Cambridge University Press.

38. Wessels, J.G.H., Gene expression during fruiting in Schizophyllum commune. *Mycological Research* 1992. 96(8): p. 609-620.

39. Kisko, K., Szilvay, G.R., Vuorimaa, E., Lemmetyinen, H., Linder, M.B., Torkkeli, M., and Serimaa, R., Self-assembled films of hydrophobin protein HFBIII from Trichoderma reesei, in *Journal of Applied Crystallography*. 2007. p. s355-s360.

40. Kisko, K., Szilvay, G.R., Vuorimaa, E., Lemmetyinen, H., Linder, M.B., Torkkeli, M., and Serimaa, R., Self-Assembled films of hydrophobin proteins HFBI and HFBII studied in Situ at the air/water interface. *Langmuir* 2009. 25(3): p. 1612-1619.

41. Lugones, L.G., Bosscher, J.S., Scholtmeyer, K., de Vries, O.M.H., and Wessels, J.G.H., An abundant hydrophobin (ABH1) forms hydrophobic rodlet layers in Agaricus bisporus fruiting bodies. *Microbiology* 1996. 142(5): p. 1321-1329.

42. Richards, W.C. and Takai, S., Novel technique for isolating microstructures present in shake cultures of the fungus ceratocystis ulmi. *Applied Microbiology* 1973. 26(3): p. 443-444.

43. Wosten, H., De Vries, O., and Wessels, J., Interfacial Self-Assembly of a Fungal Hydrophobin into a Hydrophobic Rodlet Layer. *The Plant Cell* 1993. 5(11): p. 1567-1574.

44. Wösten, H.A.B., Ruardy, T.G., van der Mei, H.C., Busscher, H.J., and Wessels, J.G.H., Interfacial self-assembly of a Schizophyllum commune hydrophobin into an insoluble amphipathic protein membrane depends on surface hydrophobicity. *Colloids and Surfaces B: Biointerfaces* 1995. 5(3): p. 189-195.

45. Wösten, H.A., Schuren, F.H., and Wessels, J.G., Interfacial self-assembly of a hydrophobin into an amphipathic protein membrane mediates fungal attachment to hydrophobic surfaces. *The EMBO Journal* 1994. 13(24): p. 5848-5854.

46. Carpenter, C.E., Mueller, R.J., Kazmierczak, P., Zhang, L., Villalon, D.K., and Van Alfen, N.K., Effect of a virus on accumulation of a tissue-specific cell-surface protein of the fungus Cryphonectria (Endothia) parasitica. *Mol Plant Microbe Interact* 1992. 5(1): p. 55-61.

47. Lugones, L.G., Wösten, H.A.B., Birkenkamp, K.U., Sjollema, K.A., Zagers, J., and Wessels, J.G.H., Hydrophobins line air channels in fruiting bodies of Schizophyllum commune and Agaricus bisporus. *Mycological Research* 1999. 103(5): p. 635-640.

48. van der Vegt, W., van der Mei, H.C., Wosten, H.A., Wessels, J.G., and Busscher, H.J., A comparison of the surface activity of the fungal hydrophobin SC3p with those of other proteins. *Biophysical Chemistry* 1996. 57: p. 253-260.

49. Lumsdon, S.O., Green, J., and Stieglitz, B., Adsorption of hydrophobin proteins at hydrophobic and hydrophilic interfaces. *Colloids and Surfaces B: Biointerfaces* 2005. 44(4): p. 172-178.

50. Cox, A.R., Cagnol, F., Russell, A.B., and Izzard, M.J., Surface properties of class II hydrophobins from Trichoderma reesei and influence on bubble stability. *Langmuir* 2007. 23: p. 7995-8002.

51. Kallio, J.M., Linder, M.B., and Rouvinen, J., Crystal structures of hydrophobin HFBII in the presence of detergent implicate the formation of fibrils and monolayer films. *The Journal of Biological Chemistry* 2007. 282(39): p. 28733-28739.

52. Alexandrov, N.A., Marinova, K.G., Gurkov, T.D., Danov, K.D., Kralchevsky, P.A., Stoyanov, S.D., Blijdenstein, T.B., Arnaudov, L.N., Pelan, E.G., and Lips, A., Interfacial layers from the protein HFBII hydrophobin: dynamic surface tension, dilatational elasticity and relaxation times. *Journal of Colloid and Interface Science* 2012. 376(1): p. 296-306.

53. Wösten, H.A.B., van Wetter, M.-A., Lugones, L.G., van der Mei, H.C., Busscher, H.J., and Wessels, J.G.H., How a fungus escapes the water to grow into the air. *Current Biology* 1999. 9(2): p. 85-88.

54. Wösten, H.A.B. and de Vocht, M.L., Hydrophobins, the fungal coat unravelled. *Biochimica et Biophysica Acta (BBA) - Reviews on Biomembranes* 2000. 1469(2): p. 79-86.

55. Wang, X., Permentier, H.P., Rink, R., Kruijtzer, J.A., Liskamp, R.M., Wosten, H.A., Poolman, B., and Robillard, G.T., Probing the self-assembly and the accompanying structural changes of hydrophobin SC3 on a hydrophobic surface by mass spectrometry. *Biophysical Journal* 2004. 87(3): p. 1919-28.

56. Askolin, S., Linder, M., Scholtmeijer, K., Tenkanen, M., Penttila, M., de Vocht, M.L., and Wosten, H.A., Interaction and comparison of a class I hydrophobin from Schizophyllum commune and class II hydrophobins from Trichoderma reesei. *Biomacromolecules* 2006. 7(4): p. 1295-301.

57. Subkowski, T., Karos, M., and Subkowski, T., Industrial performance proteins: Hydrophobin—learning from nature. *Journal of Biotechnology* 2007. 131(2, Supplement): p. S212-S213.

58. Martin, G.G., Cannon, G.C., and McCormick, C.L., Sc3p Hydrophobin Organization in Aqueous Media and Assembly onto Surfaces As Mediated by the Associated Polysaccharide Schizophyllan. *Biomacromolecules* 2000. 1(1): p. 49-60.

59. Butko, P., Buford, J.P., Goodwin, J.S., Stroud, P.A., McCormick, C.L., and Cannon, G.C., Spectroscopic evidence for amyloid-like interfacial self-assembly of hydrophobin Sc3. *Biochemical and Biophysical Research Communications* 2001. 280(1): p. 212-215.

60. de Vocht, M.L., Reviakine, I., Ulrich, W.-P., Bergsma-Schutter, W., Wosten, H.A., Vogel, H., Brisson, A., Wessels, J.G., and Robillard, G.T., Self-assembly of the hydrophobin SC3 proceeds via two structural intermediates. *Protein Science* 2002. 11(5): p. 1199-1205.

61. Ritva, S., Torkkeli, M., Paananen, A., Linder, M., Kisko, K., Knaapila, M., Ikkala, O., Vuorimaa, E., Lemmetyinen, H., and Seeck, O.H., Self-assembled structures of hydrophobins HFBI and HFBII. *Journal of Applied Crystallography* 2003. 36: p. 499-502.

62. Wang, X., Shi, F., Wosten, H.A., Hektor, H., Poolman, B., and Robillard, G.T., The SC3 hydrophobin self-assembles into a membrane with distinct mass transfer properties. *Biophysical Journal* 2005. 88(5): p. 3434-3443.

63. Fan, H., Wang, X., Zhu, J., Robillard, G.T., and Mark, A.E., Molecular dynamics simulations of the hydrophobin SC3 at a hydrophobic/hydrophilic interface. *Proteins: Structure, Function, and Bioinformatics* 2006. 64(4): p. 863-873.

64. Szilvay, G.R., Kisko, K., Serimaa, R., and Linder, M.B., The relation between solution association and surface activity of the hydrophobin HFBI from Trichoderma reesei. *FEBS Letters* 2007. 581(14): p. 2721-2726.

65. Kisko, K., Szilvay, G.R., Vainio, U., Linder, M.B., and Serimaa, R., Interactions of hydrophobin proteins in solution studied by small-angle X-ray scattering. *Biophysical Journal* 2008. 94(1): p. 198-206.

66. Rosenthal, G. and Klapp, S.H.L., Micelle and Bilayer Formation of Amphiphilic Janus Particles in a Slit-Pore. *International Journal of Molecular Sciences* 2012. 13(8): p. 9431-9446.

67. Morris, V.K., Kwan, A.H., and Sunde, M., Analysis of the Structure and Conformational States of DewA Gives Insight into the Assembly of the Fungal Hydrophobins. *Journal of Molecular Biology* 2013. 425(2): p. 244-256.

68. Israeli-Lev, G. and Livney, Y.D., Self-assembly of hydrophobin and its co-assembly with hydrophobic nutraceuticals in aqueous solutions: Towards application as delivery systems. *Food Hydrocolloids* 2014. 35(0): p. 28-35.

69. Hektor, H.J. and Scholtmeijer, K., Hydrophobins: proteins with potential. *Current Opinion in Biotechnology* 2005. 16(4): p. 434-439.

70. Sarlin, T., Nakari-Setala, T., Linder, M.B., Penttila, M., and Haikara, A., Fungal hydrophobins as predictors of the gushing activity of malt. *Journal of the Institute of Brewing* 2005. 111(2): p. 105-111.

71. Stubner, M., Lutterschmid, G., Vogel, R.F., and Niessen, L., Heterologous expression of the hydrophobin FcHyd5p from Fusarium culmorum in Pichia pastoris and evaluation of its surface activity and contribution to gushing of carbonated beverages. *Int J Food Microbiol* 2010. 141(1-2): p. 110-5.

72. Palomo, J.M., Peñas, M.M., Fernández-Lorente, G., Mateo, C., Pisabarro, A.G., Fernández-Lafuente, R., Ramírez, L., and Guisán, J.M., Solid-Phase Handling of

Hydrophobins: Immobilized Hydrophobins as a New Tool To Study Lipases. *Biomacromolecules* 2003. 4(2): p. 204-210.

73. Crilly, J.F., Russell, A.B., Cox, A.R., and Cebula, D.J., Designing Multiscale Structures for Desired Properties of Ice Cream. *Industrial & Engineering Chemistry Research* 2008. 47: p. 6362-6367.

74. Wang, Z., Wang, Y., Huang, Y., Li, S., Feng, S., Xu, H., and Qiao, M., Characterization and application of hydrophobin-dispersed multi-walled carbon nanotubes. *Carbon* 2010. 48(10): p. 2890-2898.

75. Bollschweiler, C., Karos, M., and Subkowski, T., Industrial performance proteins: An overview. *Journal of Biotechnology* 2007. 131(2, Supplement): p. S213.

76. Anon, BASF Synthesizes Hydrophobin, in *Chemical & Engineering News*. 2008, American Chemical Society. p. 17-18.

77. Wösten, H.A.B. and Scholtmeijer, K., Applications of hydrophobins: current state and perspectives. *Applied Microbiology and Biotechnology* 2015. 99(4): p. 1587-1597.

78. Arkın, H. and Janke, W., Gyration tensor based analysis of the shapes of polymer chains in an attractive spherical cage. *The Journal of Chemical Physics* 2013. 138(5): p. 054904.

79. Rudnick, J. and Gaspari, G., The Shapes of Random Walks. *Science* 1987. 237(4813): p. 384-389.

80. Benner, S.W., John, V.T., and Hall, C.K., Simulation Study of Hydrophobically Modified Chitosan as an Oil Dispersant Additive. *The Journal of Physical Chemistry B* 2015. 119(23): p. 6979-6990.

81. Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E., The Protein Data Bank. *Nucleic Acids Research* 2000. 28(1): p. 235-242.

82. Artymiuk, P.J., Blake, C.C.F., Rice, D.W., and Wilson, K.S., The structures of the monoclinic and orthorhombic forms of hen egg-white lysozyme at 6 A resolution. *Acta Crystallographica Section B* 1982. 38(3): p. 778-783.

83. Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A.E., and Berendsen, H.J.C., GROMACS: Fast, flexible, and free. *Journal of Computational Chemistry* 2005. 26(16): p. 1701-1718.

84. Jorgensen, W.L., Maxwell, D.S., and Tirado-Rives, J., Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *Journal of the American Chemical Society* 1996. 118(45): p. 11225-11236.

85. Jorgensen, W.L. and McDonald, N.A., Development of an all-atom force field for heterocycles. Properties of liquid pyridine and diazenes. *Journal of Molecular Structure: THEOCHEM* 1998. 424(1–2): p. 145-155.

86. Kaminski, G.A., Friesner, R.A., Tirado-Rives, J., and Jorgensen, W.L., Evaluation and Reparametrization of the OPLS-AA Force Field for Proteins via Comparison with Accurate Quantum Chemical Calculations on Peptides†. *The Journal of Physical Chemistry B* 2001. 105(28): p. 6474-6487.

87. McDonald, N.A. and Jorgensen, W.L., Development of an All-Atom Force Field for Heterocycles. Properties of Liquid Pyrrole, Furan, Diazoles, and Oxazoles. *The Journal of Physical Chemistry B* 1998. 102(41): p. 8049-8059.

88. Price, M.L.P., Ostrovsky, D., and Jorgensen, W.L., Gas-phase and liquid-state properties of esters, nitriles, and nitro compounds with the OPLS-AA force field. *Journal of Computational Chemistry* 2001. 22(13): p. 1340-1352.

89. Rizzo, R.C. and Jorgensen, W.L., OPLS All-Atom Model for Amines: Resolution of the Amine Hydration Problem. *Journal of the American Chemical Society* 1999. 121(20): p. 4827-4836.

90. Watkins, E.K. and Jorgensen, W.L., Perfluoroalkanes: Conformational Analysis and Liquid-State Properties from ab Initio and Monte Carlo Calculations. *The Journal of Physical Chemistry A* 2001. 105(16): p. 4118-4125.

91. Jorgensen, W.L. and Swenson, C.J., Optimized intermolecular potential functions for amides and peptides. Structure and properties of liquid amides. *Journal of the American Chemical Society* 1985. 107(3): p. 569-578.

92. Ponder, J.W. and Case, D.A., Force fields for protein simulations. *Advances in Protein Chemistry* 2003. 66: p. 27-85.

93. Berendsen, H.J.C., Grigera, J.R., and Straatsma, T.P., The missing term in effective pair potentials. *The Journal of Physical Chemistry* 1987. 91(24): p. 6269-6271.

94. Vácha, R., Siu, S.W.I., Petrov, M., Böckmann, R.A., Barucha-Kraszewska, J., Jurkiewicz, P., Hof, M., Berkowitz, M.L., and Jungwirth, P., Effects of Alkali Cations and Halide Anions on the DOPC Lipid Membrane†. *The Journal of Physical Chemistry A* 2009. 113(26): p. 7235-7243.

95. Liyana-Arachchi, T.P., Zhang, Z., Ehrenhauser, F.S., Avij, P., Valsaraj, K.T., and Hung, F.R., Bubble bursting as an aerosol generation mechanism during an oil spill in the deep-sea environment: molecular dynamics simulations of oil alkanes and dispersants in atmospheric air/salt water interfaces. *Environmental Science: Processes & Impacts* 2014. 16(1): p. 53-64.

96. Liyana-Arachchi, T.P., Valsaraj, K., and Hung, F.R., Ice growth from supercooled aqueous solutions of reactive oxygen species. *Theoretical Chemistry Accounts* 2012. 132(1): p. 1-13.

97. Auffinger, P., Cheatham, T.E., and Vaiana, A.C., Spontaneous Formation of KCl Aggregates in Biomolecular Simulations: A Force Field Issue? *Journal of Chemical Theory and Computation* 2007. 3(5): p. 1851-1859.

98. Patriksson, A., Marklund, E.G., and van Der Spoel, D., Protein structures under electrospray conditions. *Biochemistry* 2007. 46(4): p. 933-945.

99. Bussi, G., Donadio, D., and Parrinello, M., Canonical sampling through velocity rescaling. *Journal of Chemical Physics* 2007. 126(1).

100. Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., DiNola, A., and Haak, J.R., Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics* 1984. 81(8): p. 3684-3690.

101. Miyamoto, S. and Kollman, P.A., Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *Journal of Computational Chemistry* 1992. 13(8): p. 952-962.

102. Hess, B., Bekker, H., Berendsen, H.J.C., and Fraaije, J.G.E.M., LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry* 1997. 18(12): p. 1463-1472.

103. Darden, T., York, D., and Pedersen, L., Particle mesh Ewald: An N · log(N) method for Ewald sums in large systems. *The Journal of Chemical Physics* 1993. 98(12): p. 10089-10092.

104. Barker, J.A. and Watts, R.O., Monte Carlo studies of the dielectric properties of water-like models. *Molecular Physics* 1973. 26(3): p. 789-792.

105. Watts, R.O., Monte Carlo studies of liquid water. *Molecular Physics* 1974. 28(4): p. 1069-1083.

106. Dang, L.X., Computational Study of Ion Binding to the Liquid Interface of Water†. *The Journal of Physical Chemistry B* 2002. 106(40): p. 10388-10394.

107. Miller, C.A., Abbott, N.L., and de Pablo, J.J., Surface Activity of Amphiphilic Helical β-Peptides from Molecular Dynamics Simulation. *Langmuir* 2009. 25(5): p. 2811-2823.

108. Trzesniak, D., Kunz, A.-P.E., and van Gunsteren, W.F., A Comparison of Methods to Compute the Potential of Mean Force. *ChemPhysChem* 2007. 8(1): p. 162-169.

109. Liyana-Arachchi, T.P., Hansel, A.K., Stevens, C., Ehrenhauser, F.S., Valsaraj, K.T., and Hung, F.R., Molecular Modeling of the Green Leaf Volatile Methyl Salicylate on

Atmospheric Air/Water Interfaces. *The Journal of Physical Chemistry A* 2013. 117(21): p. 4436-4443.

110. Liyana-Arachchi, T.P., Stevens, C., Hansel, A.K., Ehrenhauser, F.S., Valsaraj, K.T., and Hung, F.R., Molecular simulations of green leaf volatiles and atmospheric oxidants on air/water interfaces. *Physical Chemistry Chemical Physics* 2013. 15(10): p. 3583-3592.

111. Liyana-Arachchi, T.P., Valsaraj, K.T., and Hung, F.R., Molecular Simulation Study of the Adsorption of Naphthalene and Ozone on Atmospheric Air/Ice Interfaces. *The Journal of Physical Chemistry A* 2011. 115(33): p. 9226-9236.

112. Liyana-Arachchi, T.P., Valsaraj, K.T., and Hung, F.R., Ice Growth from Supercooled Aqueous Solutions of Benzene, Naphthalene, and Phenanthrene. *The Journal of Physical Chemistry A* 2012. 116(33): p. 8539-8546.

113. Liyana-Arachchi, T.P., Valsaraj, K.T., and Hung, F.R., Adsorption of Naphthalene and Ozone on Atmospheric Air/Ice Interfaces Coated with Surfactants: A Molecular Simulation Study. *The Journal of Physical Chemistry A* 2012. 116(10): p. 2519-2528.

114. Marrink, S.J., de Vries, A.H., and Mark, A.E., Coarse grained model for semiquantitative lipid simulations. *Journal of Physical Chemistry B* 2004. 108(2): p. 750-760.

115. Marrink, S.J., Risselada, H.J., Yefimov, S., Tieleman, D.P., and de Vries, A.H., The MARTINI Force Field: coarse grained model for biomolecular simulations. *Journal of Physical Chemistry B* 2007. 111: p. 7812-7824.

116. Monticelli, L., Kandasamy, S.K., Periole, X., Larson, R.G., Tieleman, D.P., and Marrink, S.J., The MARTINI coarse-grained force field: Extension to proteins. *Journal of Chemical Theory and Computation* 2008. 4(5): p. 819-834.

117. de Jong, D.H., Singh, G., Bennett, W.F.D., Arnarez, C., Wassenaar, T.A., Schäfer, L.V., Periole, X., Tieleman, D.P., and Marrink, S.J., Improved Parameters for the Martini Coarse-Grained Protein Force Field. *Journal of Chemical Theory and Computation* 2012. 9(1): p. 687-697.

118. Yesylevskyy, S.O., Schäfer, L.V., Sengupta, D., and Marrink, S.J., Polarizable Water Model for the Coarse-Grained MARTINI Force Field. *PLOS Computational Biology* 2010. 6(6): p. e1000810.

119. Periole, X., Cavalli, M., Marrink, S.J., and Ceruso, M.A., Combining an Elastic Network With a Coarse-Grained Molecular Force Field: Structure, Dynamics, and Intermolecular Recognition. *Journal of Chemical Theory and Computation* 2009. 5(9): p. 2531-2543.

120. Tirion, M.M., Large Amplitude Elastic Motions in Proteins from a Single-Parameter, Atomic Analysis. *Physical Review Letters* 1996. 77(9): p. 1905-1908.

121. Doruker, P., Atilgan, A.R., and Bahar, I., Dynamics of proteins predicted by molecular dynamics simulations and analytical approaches: application to alpha-amylase inhibitor. *Proteins* 2000. 40(3): p. 512-24.

122. Atilgan, A.R., Durell, S.R., Jernigan, R.L., Demirel, M.C., Keskin, O., and Bahar, I., Anisotropy of fluctuation dynamics of proteins with an elastic network model. *Biophysical Journal* 2001. 80(1): p. 505-15.

123. Kundu, S., Melton, J.S., Sorensen, D.C., and Phillips Jr, G.N., Dynamics of Proteins in Crystals: Comparison of Experiment with Simple Models. *Biophysical Journal* 2002. 83(2): p. 723-732.

124. Sen, T.Z., Feng, Y., Garcia, J.V., Kloczkowski, A., and Jernigan, R.L., The Extent of Cooperativity of Protein Motions Observed with Elastic Network Models Is Similar for Atomic and Coarser-Grained Models. *Journal of Chemical Theory and Computation* 2006. 2(3): p. 696-704.

125. Kondrashov, D.A., Van Wynsberghe, A.W., Bannen, R.M., Cui, Q., and Phillips Jr, George N., Protein Structural Variation in Computational Models and Crystallographic Data. *Structure* 2007. 15(2): p. 169-177.

126. Zheng, W., Anharmonic Normal Mode Analysis of Elastic Network Model Improves the Modeling of Atomic Fluctuations in Protein Crystal Structures. *Biophysical Journal* 2010. 98(12): p. 3025-3034.

127. Marrink, S.J., Risselada, H.J., Yefimov, S., Tieleman, D.P., and de Vries, A.H., The MARTINI force field: Coarse grained model for biomolecular simulations. *Journal of Physical Chemistry B* 2007. 111(27): p. 7812-7824.

128. Marrink, S.J. and Tieleman, D.P., Perspective on the Martini model. *Chemical Society Reviews* 2013. 42(16): p. 6801-6822.

129. Alvarez, E., Rendo, R., Sanjurjo, B., Sánchez-Vilas, M., and Navaza, J.M., Surface Tension of Binary Mixtures of Water + N-Methyldiethanolamine and Ternary Mixtures of This Amine and Water with Monoethanolamine, Diethanolamine, and 2-Amino-2-methyl-1-propanol from 25 to 50 °C. *Journal of Chemical & Engineering Data* 1998. 43(6): p. 1027-1029.

130. Jasper, J.J., The Surface Tension of Pure Liquid Compounds. *Journal of Physical and Chemical Reference Data* 1972. 1(4): p. 841-1010.

131. Zeppieri, S., Rodríguez, J., and López de Ramos, A.L., Interfacial Tension of Alkane + Water Systems†. *Journal of Chemical & Engineering Data* 2001. 46(5): p. 1086-1088.

132. Huston, K.J. and Larson, R.G., Reversible and Irreversible Adsorption Energetics of Poly(ethylene glycol) and Sorbitan Poly(ethoxylate) at a Water/Alkane Interface. *Langmuir* 2015. 31(27): p. 7503-7511.

133. Reichert, M.D. and Walker, L.M., Interfacial Tension Dynamics, Interfacial Mechanics, and Response to Rapid Dilution of Bulk Surfactant of a Model Oil–Water-Dispersant System. *Langmuir* 2013. 29(6): p. 1857-1867.

134. Houmadi, S., Ciuchi, F., Santo, M.P.D., Stefano, L.D., Rea, I., Giardina, P., Armenante, A., Lacaze, E., and Giocondo, M., Langmuir-Blodgett Film of Hydrophobin Protein from Pleurotus ostreatus at the Air-Water Interface. *Langmuir* 2008. 24: p. 12953-12957.

135. Lobanov, M.Y., Bogatyreva, N.S., and Galzitskaya, O.V., Radius of gyration as an indicator of protein structure compactness. *Molecular Biology* 2008. 42(4): p. 623-628.

136. Marklund, E.G., Larsson, D.S.D., van der Spoel, D., Patriksson, A., and Caleman, C., Structural stability of electrosprayed proteins: temperature and hydration effects. *Physical Chemistry Chemical Physics* 2009. 11(36): p. 8069-8078.

137. Pettit, F.K., Bare, E., Tsai, A., and Bowie, J.U., HotPatch: A Statistical A pproach to Finding Biologically Relevant Features on Protein Surfaces. *Journal of Molecular Biology* 2007. 369(3): p. 863-879.

138. Lee, B. and Richards, F.M., The interpretation of protein structures: Estimation of static accessibility. *Journal of Molecular Biology* 1971. 55(3): p. 379-400.

139. Miller, S., Lesk, A.M., Janin, J., and Chothia, C., The accessible surface area and stability of oligomeric proteins. *Nature* 1987. 328(6133): p. 834-836.

140. Shrake, A. and Rupley, J.A., Environment and exposure to solvent of protein atoms. Lysozyme and insulin. *Journal of Molecular Biology* 1973. 79(2): p. 351-371.

141. Marsh, Joseph A. and Teichmann, Sarah A., Relative Solvent Accessible Surface Area Predicts Protein Conformational Changes upon Binding. *Structure* 2011. 19(6): p. 859-867.

142. Lins, L., Thomas, A., and Brasseur, R., Analysis of accessible surface of residues in proteins. *Protein Science* 2003. 12(7): p. 1406-1417.

143. Chang, M., Choi, D., Fu, B., and Reichmanis, E., Solvent Based Hydrogen Bonding: Impact on Poly(3-hexylthiophene) Nanoscale Morphology and Charge Transport Characteristics. *ACS Nano* 2013. 7(6): p. 5402-5413.

144. Schmidtke, S., Russo, P., Nakamatsu, J., Buyuktanir, E., Turfan, B., Temyanko, E., and Negulescu, I., Thermoreversible Gelation of Isotropic and Liquid Crystalline Solutions of a “Sticky” Rodlike Polymer. *Macromolecules* 2000. 33(12): p. 4427-4432.

145. Poche, D.S., Daly, W.H., and Russo, P.S., Synthesis and Some Solution Properties of Poly(.gamma.-stearyl .alpha.,L-glutamate). *Macromolecules* 1995. 28(20): p. 6745-6753.

146. Balamurugan, S.S., Soto-Cantu, E., Cueto, R., and Russo, P.S., Preparation of Organosoluble Silica−Polypeptide Particles by "Click" Chemistry. *Macromolecules* 2010. 43(1): p. 62-70.

147. Jorgensen, W.L. and Tirado-Rives, J., Potential energy functions for atomic-level simulations of water and organic and biomolecular systems. *Proceedings of the National Academy of Sciences of the United States of America* 2005. 102(19): p. 6665-6670.

148. Caleman, C., van Maaren, P.J., Hong, M., Hub, J.S., Costa, L.T., and van der Spoel, D., Force Field Benchmark of Organic Liquids: Density, Enthalpy of Vaporization, Heat Capacities, Surface Tension, Isothermal Compressibility, Volumetric Expansion Coefficient, and Dielectric Constant. *Journal of Chemical Theory and Computation* 2012. 8(1): p. 61-74.

149. Siu, S.W.I., Pluhackova, K., and Böckmann, R.A., Optimization of the OPLS-AA Force Field for Long Hydrocarbons. *Journal of Chemical Theory and Computation* 2012. 8(4): p. 1459-1470.

150. Shirts, M.R., Pitera, J.W., Swope, W.C., and Pande, V.S., Extremely precise free energy calculations of amino acid side chain analogs: Comparison of common molecular mechanics force fields for proteins. *The Journal of Chemical Physics* 2003. 119(11): p. 5740-5761.

151. Hub, J.S., Caleman, C., and van der Spoel, D., Organic molecules on the surface of water droplets - an energetic perspective. *Physical Chemistry Chemical Physics* 2012. 14(27): p. 9537-9545.

152. Bennett, C.H., Efficient estimation of free energy differences from Monte Carlo data. *Journal of Computational Physics* 1976. 22(2): p. 245-268.

153. Liyana-Arachchi, T.P., Zhang, Z., Vempati, H., Hansel, A.K., Stevens, C., Pham, A.T., Ehrenhauser, F.S., Valsaraj, K.T., and Hung, F.R., Green Leaf Volatiles on Atmospheric Air/Water Interfaces: A Combined Experimental and Molecular Simulation Study. *Journal of Chemical & Engineering Data* 2014. 59(10): p. 3025-3035.

154. Letcher, T.M. and Domańska, U., The excess volumes of (tributylamine + an ether) at the temperature 298.15 K. *The Journal of Chemical Thermodynamics* 1994. 26(11): p. 1241-1247.

155. Hossenlopp, I.A. and Scott, D.W., Vapor heat capacities and enthalpies of vaporization of six organic compounds. *The Journal of Chemical Thermodynamics* 1981. 13(5): p. 405-414.

156. V.I., M., Thermodynamic Properties of Tetrahydrofuran in the Range 8 - 322 K. *Viniti* 1978: p. 1-21.

157. Dorofeeva, O.V., Ideal gas thermodynamic properties of oxygen heterocyclic compounds Part 1. Three-membered, four-membered and five-membered rings. *Thermochimica Acta* 1992. 194: p. 9-46.

158. Hawkins, G.D., Cramer, C.J., and Truhlar, D.G., Universal Quantum Mechanical Model for Solvation Free Energies Based on Gas-Phase Geometries. *The Journal of Physical Chemistry B* 1998. 102(17): p. 3257-3271.

159. Gonçalves, P.F.B. and Stassen, H., Calculation of the free energy of solvation from molecular dynamics simulations. *Pure and Applied Chemistry* 2004. 76(1): p. 231-240.

160. Bhatta, R.S., Yimer, Y.Y., Perry, D.S., and Tsige, M., Improved Force Field for Molecular Modeling of Poly(3-hexylthiophene). *The Journal of Physical Chemistry B* 2013. 117(34): p. 10035-10045.

161. Marcon, V. and Raos, G., Free Energies of Molecular Crystal Surfaces by Computer Simulation: Application to Tetrathiophene. *Journal of the American Chemical Society* 2006. 128(5): p. 1408-1409.

162. Bhatta, R.S., Yimer, Y.Y., Tsige, M., and Perry, D.S., Conformations and torsional potentials of poly(3-hexylthiophene) oligomers: Density functional calculations up to the dodecamer. *Computational and Theoretical Chemistry* 2012. 995: p. 36-42.

163. Groot, R.D. and Warren, P.B., Dissipative particle dynamics: Bridging the gap between atomistic and mesoscopic simulation. *The Journal of Chemical Physics* 1997. 107(11): p. 4423-4435.

164. Murtola, T., Bunker, A., Vattulainen, I., Deserno, M., and Karttunen, M., Multiscale modeling of emergent materials: biological and soft matter. *Physical Chemistry Chemical Physics* 2009. 11(12): p. 1869-1892.

165. Rekvig, L., Kranenburg, M., Vreede, J., Hafskjold, B., and Smit, B., Investigation of Surfactant Efficiency Using Dissipative Particle Dynamics. *Langmuir* 2003. 19(20): p. 8195-8205.

166. Plimpton, S., Fast Parallel Algorithms for Short-Range Molecular Dynamics. *Journal of Computational Physics* 1995. 117(1): p. 1-19.

167. Kremer, K. and Grest, G.S., Dynamics of entangled linear polymer melts: A molecular - dynamics simulation. *The Journal of Chemical Physics* 1990. 92(8): p. 5057-5086.

Vita

Yuwu Chen, a Han Chinese from Hunan Province, People's Republic of China, received his first Master's Degree in Environmental Engineering from Nanjing University of Science and Technology (NJUST), second Master's Degree in Chemical & Environmental Engineering from the University of Toledo (UT) and third Master's Degree in Applied Statistics from Louisiana State University (LSU). He was accepted into the LSU School of Engineering majoring in chemical engineering studies. He anticipates graduating with his Ph.D. degree in May 2017. He plans to continue promoting LSU for many years to come.